General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

Produced by the NASA Center for Aerospace Information (CASI)

PROCEEDINGS

CALTECH/JPL CONFERENCE ON IMAGE PROCESSING TECHNOLOGY, DATA SOURCES AND SOFTWARE FOR COMMERCIAL AND SCIENTIFIC APPLICATIONS

CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

November 3–5, 1976

1. Report No. 43-30	2. Government Accession No.	3. Recipient's Catalog No.

4. Title and Subtitle	5. Report Date
PROCEEDINGS, CALTECH/JPL CONFERENCE ON IMAGE PROCESSING TECHNOLOGY, DATA SOURCES AND SOFTWARE FOR COMMERCIAL AND SCIENTIFIC APPLICATIONS	November 3-5, 1976
	6. Performing Organization Code

7. Author(s) Gunther H. Redmann, Editor	8. Performing Organization Report No.

9. Performing Organization Name and Address	10. Work Unit No.
JET PROPULSION LABORATORY California Institute of Technology 4800 Oak Grove Drive Pasadena, California 91103	11. Contract or Grant No. NAS 7-100
	13. Type of Report and Period Covered Special Publication

12. Sponsoring Agency Name and Address	
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION Washington, D.C. 20546	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

The Caltech/JPL Conference on Image Processing Technology, Data Sources and Software for Commercial and Scientific Applications was held at the California Institute of Technology, Pasadena, California, November 3-5, 1976, under the sponsorship of the Jet Propulsion Laboratory, the California Institute of Technology, and the NASA Office of Applications. The purpose of the conference was to disseminate information about recent advances in image processing technology and to describe new applications to the user community in order to stimulate the development and transfer of this technology to industrial and commercial applications. The Proceedings contains 37 papers and abstracts, including many illustrations (some in color) and provides a single reference source for the user community regarding the ordering and obtaining of NASA-developed image-processing software and science data.

17. Key Words (Selected by Author(s))	18. Distribution Statement
Geosciences (General) Earth Resources Mathematical and Computer Sciences (General) Urban Technology and Transportation	Unclassified -- Unlimited

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	10	

PROCEEDINGS

CALTECH/JPL CONFERENCE ON IMAGE PROCESSING TECHNOLOGY, DATA SOURCES AND SOFTWARE FOR COMMERCIAL AND SCIENTIFIC APPLICATIONS

CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA

November 3–5, 1976

CONFERENCE COMMITTEE

William B. Green, JPL, Chairman

Gunther H. Redmann, JPL, Conference Coordinator
Fred C. Billingsley, NASA Office of Applications/JPL
Robert B. Gilmore, California Institute of Technology

CHAIRMAN'S WELCOME

It is a pleasure to welcome you to the Caltech/JPL Conference on Image Processing Technology, Data Sources and Software for Commercial and Scientific Applications. The Landsat imagery, the striking images of Venus and Mercury, and the recent dramatic images of Mars returned by Viking have served to heighten a growing interest in computer processing of imagery for a variety of applications both in space exploration and here on earth. Digital image processing is an important tool in helping man understand the complex universe that surrounds him and aids significantly in increasing man's understanding of his own environment and the resources available to him. Each of you here at this conference may ultimately play an important role in the expansion of man's conception of his universe, and it is my hope that the techniques and tools we discuss at this conference will aid you in fulfilling your own role as part of this dramatic story.

We have worked hard to make this conference serve as an effective means for disseminating recent advances in image processing technology to a diverse audience. In particular, the technical papers present previously unpublished techniques that have been developed within JPL's Image Processing Laboratory during the past two years and that have potential application beyond the areas described here.
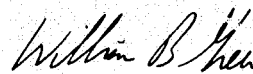
In addition to the JPL participants, we are fortunate to have with us the directors of three organizations concerned with the distribution of imagery and computer software produced by NASA, who will provide details of the resources available to you through their respective centers.

We have designed three panel discussions to serve as forums for discussion and analysis of problems of contemporary interest. The three panels include representatives from the private sector, the university community, and several NASA centers currently involved in projects that are highly user-oriented. I am extremely pleased that a large number of highly knowledgeable individuals from across the country have accepted our invitation to participate with us.

We have selected a dynamic and talented individual, Gene Thorley, to serve as our keynote speaker, and I am sure his talk will reflect the expanding horizons of digital image processing.

Finally, on a personal note, there are several individuals I wish to thank for their assistance with this conference. First, I extend my personal thanks to the members of the Conference Committee: Gunther Redmann of JPL, Fred Billingsley (currently on assignment from JPL to NASA Headquarters), and Robert Gilmore of Caltech. Next, I wish to thank two dedicated secretaries, Carol Snyder and Pam Houston, who both possess enormous quantities of patience, fortitude and skill and who were indispensable during the last few hectic months. Finally, I wish to thank Gunther Redmann, the Conference Coordinator, for his vision, his dedication, and his ability to keep track of every detail concerning this conference. Without his efforts, particularly during the months when Viking occupied a large portion of my own time, this conference would never have come to fruition.

Again, welcome to Pasadena, and I look forward with you to an exciting conference.

William B. Green
Chairman

CONTENTS

APPLICATIONS OF IMAGE PROCESSING

Gene A. Thorley

EROS PROGRAM OFFICE

The recent landing of the Viking spacecraft on Mars was a remarkable scientific and engineering achievement. The pictures of the inhospitable Martian surface serve to remind us, however, of the importance of maintaining our fragile life conditions here on Earth. Recent advances in remote sensing technology, coupled with improved remote sensing platforms, offer great promise for providing data of environmental parametrs worldwide. To a large extent, the amount of usable information obtained from these data will be dependent on continued improvement in image processing technology.

Image processing is a broad science, ranging from the production of photo like images from electronic data, to statistical and optical classification of data using spectral and spatial information.

In a like way, the applications of image processing are varied, with uses in medicine, earth resources survey, national defense, meteorology, and space science, to name but a few.

Applications of image processing in earth resources survey have been particularly impressive, with successful uses in vegetation inventory, water management, mineral and petroleum exploration and urban development. If we are to effectively utilize the data from existing and planned remote sensing systems, however, more attention must be given to meeting user needs in a cost-effective manner, determining the tradeoffs of spatial resolution, spectral diversity and temporal frequency, and overcoming institutional problems which impede operational acceptance.

NEW TECHNIQUES FOR DIGITAL IMAGE ENHANCEMENT

A. A. Schwartz

Jet Propulsion Laboratory
Pasadena, California

ABSTRACT

One of the more widespread applications of
digital image processing is the application of
image enhancement algorithms. Basically, enhance-
ment processes are designed to increase image
detail and allow more effective interpretation of
the images. Contrast enhancement is a technique
whereby the intensity distribution of a digital
image is modified in order to utilize full dynamic
range of the output display device. A number of
contrast enhancement techniques, developed for
enhancement of spaceflight photography at JPL's
Image Processing Laboratory are described in this
paper. All algorithms described in this paper are
automated, in that the algorithms perform histogram
analysis to produce images with flat, Gaussian or
parameter controlled characteristics.

INTRODUCTION

Enhancement of images is one of the classes
of tasks routinely performed at most installations
doing digital image processing. At the Image
Processing Laboratory (IPL) at JPL, determination
of camera signature effects (calibration) and
removal of these effects from images (decalibration)
is another major task. Decalibration might be con-
sidered image enhancement, since it improves visual
interpretation of the resultant image. However,
the techniques to be discussed here are of a more
subjective nature, although they do have applica-
tions in compensating for some of the effects of
instrument signature.

The contrast enhancement described in this
paper was developed at the Image Processing Labora-
tory to support the requirements of a variety of
unmanned planetary missions. It is our purpose to
give a detailed description of the algorithms, the
implementation considerations, and the rationale
used in applying each process to a particular image.

BASIC THEORY OF CONTRAST ENHANCEMENT

The general goal of contrast enhancement is
to produce an image that optimally utilizes the
full dynamic range of a selected playback device.
The device may be a television screen, photo-
graphic film, or some other more exotic piece of
equipment whose utlimate purpose is to present an
image for visual interpretation. We assume that
the device is well calibrated and that we only
need consider the characteristics of the image in
question.

First, let us consider the conditions that
make image contrast enhancement desirable. Many
factors enter into the intrinsic contrast charac-
teristics of a recorded image. These factors
include camera exposure settings, possible
vignetting effects, vidicon sensitivity and spa-
tially dependent radiometric response character-
istics (called shading), planetary albedo, solar
lighting effects, etc. Quantitative definition of
some of these factors may be unavailable, and the

factors may even vary with time. Since it is not
usually possible to quantify all of these factors,
a recorded image often does not span the dynamic
range to which it is digitized. Often processing
of the image (such as removal of camera induced
shading effects) further reduces the dynamic range.

The histogram may be used to determine the
contrast characteristics of an image and aid in
performing enhancement. A histogram of a digital
image is a discrete function which describes the
frequency distribution of the DN (digital number
or gray level) values of the picture elements.
The digital values present in a particular imaging
system are represented by

$$F = \{f_d : d = 0, 1, \ldots, m_d\} \tag{1}$$

where m_d represents the maximum possible digitized
DN value of the recording device. Then $F(k)$ is the
frequency of occurrence of DN value k in the image.

Often the high and low DN values of the digi-
tized image represent saturation effects due to
exposure problems or other factors. Thus it is
sometimes useful to define a modified histogram,
some of whose function values have been set to
zero. In this manner, we may exclude from con-
sideration some of the DN values of the original
image, obtaining

$$F' = \{f'_d : f'_d = f_d \text{ or } f'_d = 0 \text{ for selected d}\} \tag{2}$$

For notational convenience, we will now omit the
prime, implying that all enhancement algorithms
described in this paper assume that exclusion of
some DN values may have occurred prior to histo-
gram analysis based on a priori knowledge of
factors specific to the digital system employed.

One type of contrast enhancement may be
obtained by finding a piecewise linear transforma-
tion which causes user specified percentages of
the picture elements to be greater than (less than)
certain DN values in the histogram of the enhanced
images denoted as H and L. Figure 1 illustrates
this type of transformation as defined by

$$g_d = (f_d - f_e) \cdot \frac{(H-L)}{(f_h - f_e)} + L \tag{3}$$

with constraints

$$\sum_{r=h+1}^{m_d} f_r \geq P_H \quad , \quad \sum_{r=0}^{\ell-1} f_r \geq P_L$$

In practice it is desirable to replace H by m_d and
L by zero in Eq. (3), but maintain the constraints.
Note that this is not equivalent to redefining H
and L; P_H still specifies a perdentage of DN values

that are to be greater than H in the resultant
histogram and similarly for L. The linear enhance-
ment described above works well for images with
smooth histograms with a single maximum. Often,
however, this is not the case, and algorithms
employing more complex analyses of the input histo-
gram and non-linear output intensity distributions
are employed.

Sometimes a histogram will exhibit a bimodal
or multimodal distribution. Figure 2 illustrates
an image with such a distribution. This is a
Mariner 9 image of the south polar cap of Mars,
after contrast enhancement performed using the
linear stretch technique. There are channels of
dirt and ice interleaved in the bright polar cap.
For this type of image the linear transformation
is clearly inadequate.

CONTRAST ENHANCEMENT BASED ON CUMULATIVE DISTRIBUTION FUNCTION TRANSFORMATIONS

Another type of contrast enhancement, some-
times called cumulative distribution function (CDF)
matching may be obtained as follows. First, we
specify a desired histogram frequency distribution
B (e.g., a Gaussian distribution). Now we find a
transformation $t:F \rightarrow G$ such that the resultant
histogram matches as closely as possible the
desired distribution. Figure 3 illustrates the
effect of two such transformations on the unen-
hanced version of the same image depicted in Fig.
2.

In Fig. 3a, the desired transformation was to
a Gaussian defined so that the mean is centered at
mid-range of the histogram and the standard dis-
tribution of the output (Gaussian) histogram con-
tains the entire DN range. In Fig. 3b, the output
histogram has also been specified as a Gaussian
distribution, but the standard deviation of the
output distribution is one-sixth on the entire DN
range.

The uniform output DN distribution is another
type of idealized function used to generate a con-
trast transformation by the CDF matching method.
The uniform output distribution is a function all
of whose values are equal. This type of function
would at first seem to be the desired result of a
transformation designed to optimize image contrast.
However, in some cases a uniform output distribu-
tion proves too harsh, and the Gaussian or linear
transformation provides a better result. Figure 4
is an example of the uniform CDF matching trans-
formation.

A third function used in CDF matching trans-
formation is a power law function. This function
is more properly described as a functional, since
it is the generator for an entire class of func-
tions. The power law is more flexible than either
the Gaussian or uniform distributions since it can
be used to generate a wide variety of distributions.

The power law generator is defined by $y = 1 - |x|^n$, on the closed interval $(-1,1)$. We then
define G by

$$G(\ell) = y(-1)$$

$$G(k) = y\left[k \frac{2}{h-\ell} - 1\right]$$

$$G(h) = y(1)$$

all other $G = 0$

Figure 5 illustrates the graph of this func-
tion for various values of n.

One of the interesting properties of the power
law functional is that for all finite values of n,
$G(\ell) = G(h) = 0$. Thus we may construct a CDF
matching distribution which is arbitrarily close to
being uniform, but which does not cause any DN
values to be assigned to the high and low ends of
the distribution. This property is a consequence
of the fact that the power law functional generates
a sequence of functions which is convergent, but
not uniformly convergent.

Note that the CDF matching technique may also
be used to transform an image so that the histogram
of the resultant image resembles as closely as
possible the histogram of some other image.

Figure 6 indicates several versions of an
image of the ocean floor in the "famous" region of
the mid-Atlantic ridge. The original version,
Fig. 6a, is the result of a digital scan of the
original 70mm negative. Figure 6b is a Gaussian
contrast enhancement, and Fig. 6c shows the result
of a uniform output distribution CDF matching.
Fig. 6d shows the utility of contrast enhancement
combined with another subjective enhancement tech-
nique, high pass filtering. The input image has
been filtered digitally, using a filter designed to
enhance local scene detail and suppress low fre-
quency (slowly varying) intensity variations. A
Gaussian stretch has been applied to the filtered
version, providing an optimum display of the high
frequency scene components that is matched to the
full dynamic range of the output display medium.

ACKNOWLEDGEMENT
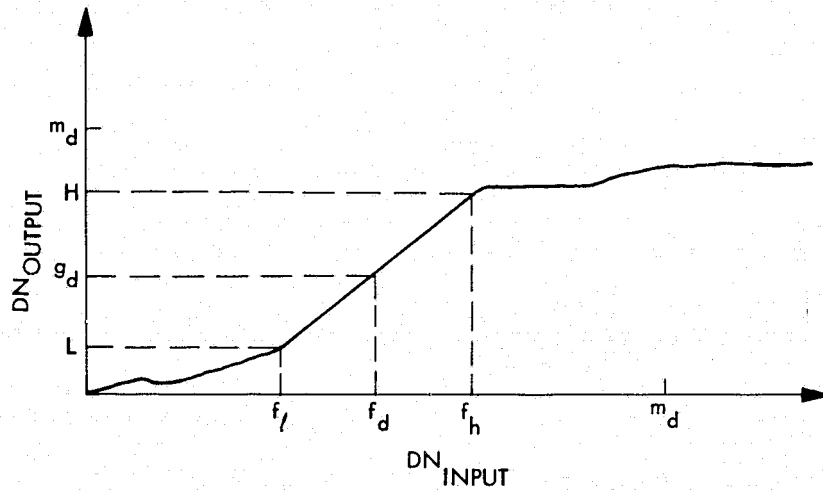
$$g_d{}' = (f_d - f_l) \frac{(H-L)}{(f_h - f_l)} \times L$$

CONSTRAINTS ARE

$$\sum_{r=h+1}^{m_d} f_r \geq P_H$$

$$\sum_{r=o}^{l-1} f_r \geq P_L$$

P_H AND P_L ARE PARAMETERS

Figure 1. Illustration of linear contrast enhancement algorithm.
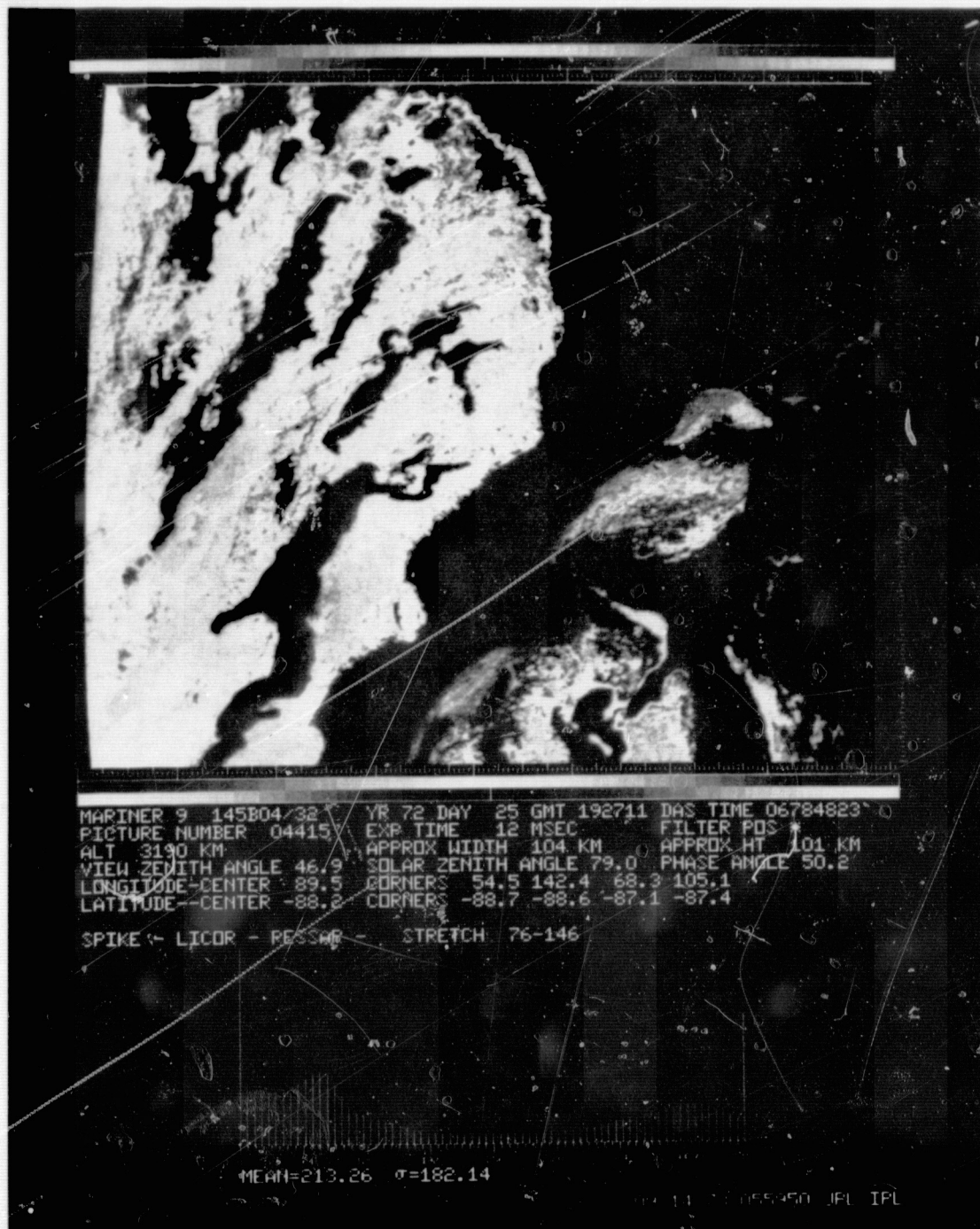
Figure 2. South polar cap region of Mars photographed by Mariner 9. Linear
contrast enhancement has been performed.

Figure 3a. Application of Gaussian stretch to the same image as Fig. 2. The output distribution function has been designed such that the standard distribution of the output distribution spans the full histogram.
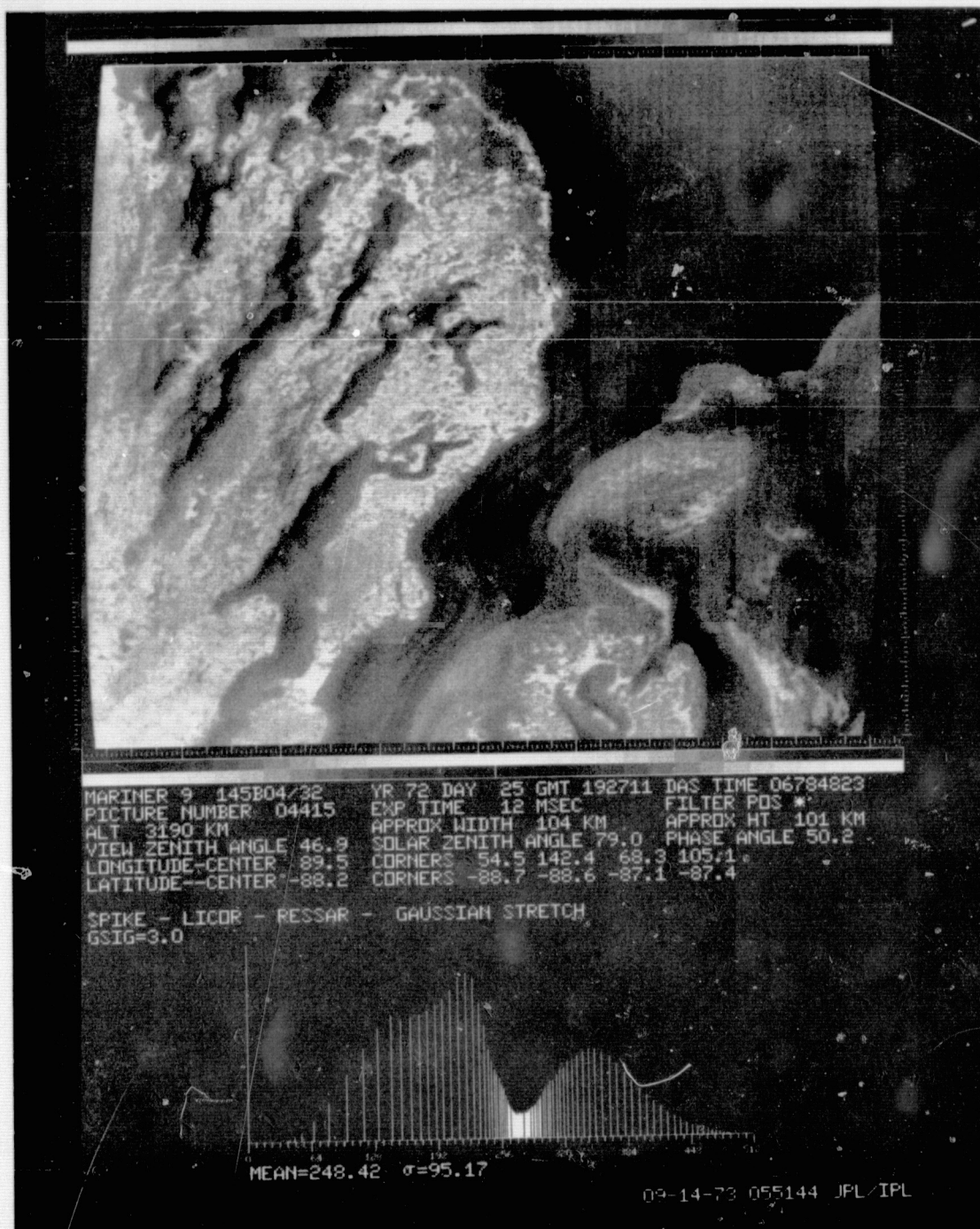
Figure 3b. Application of Gaussian stretch to the same image as Fig. 2. The output distribution function has been designed to have a standard deviation equal to one-sixth of the histogram range.
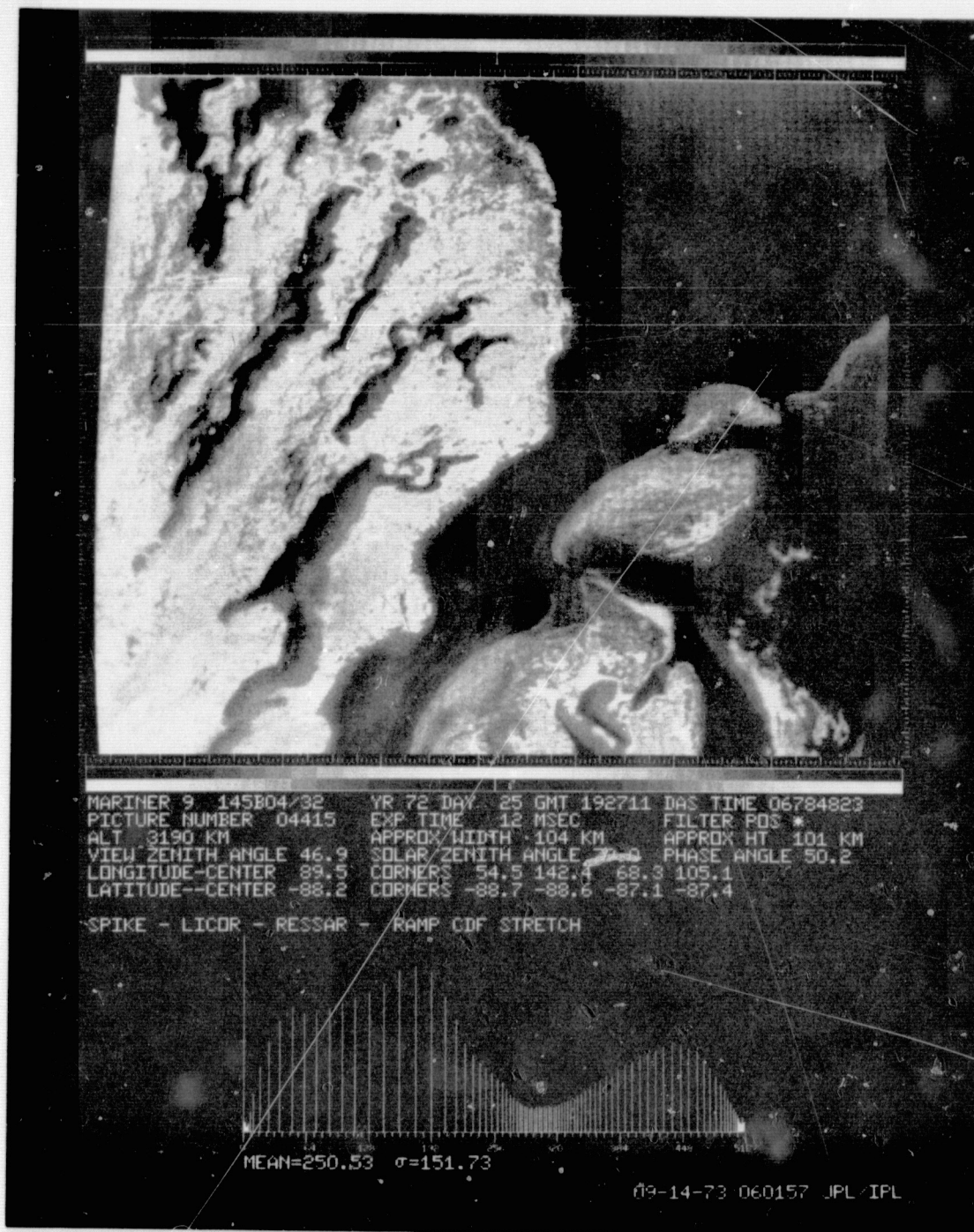
Figure 4. Application of stretch designed to achieve uniform output distribution function. Note result is more harsh than result shown in Fig. 3b.
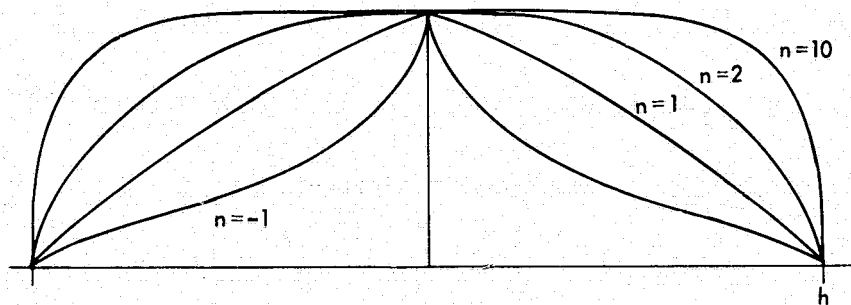
Figure 5. Power law generator function for n = -1, 1, 2 and 10.
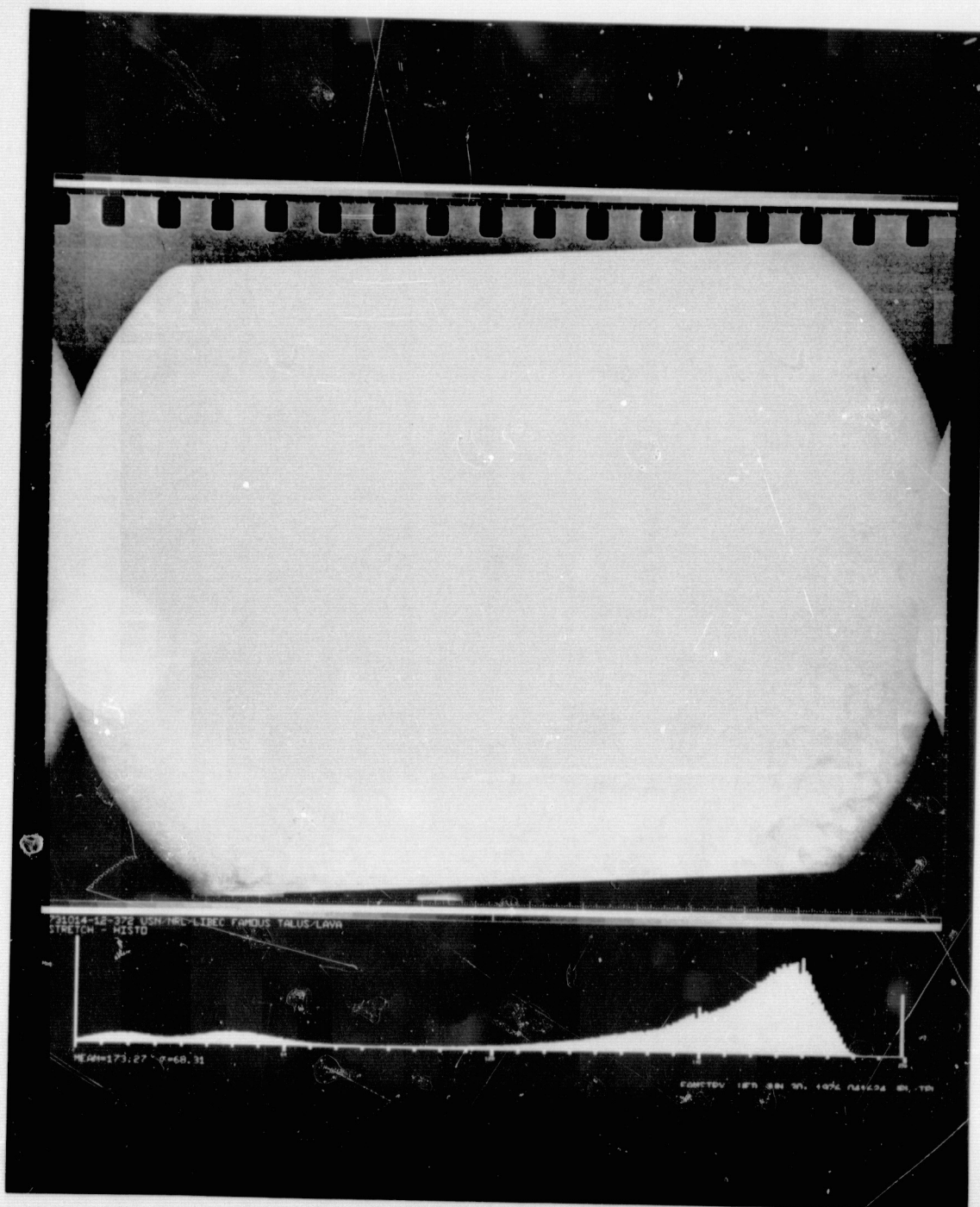
Figure 6a.   Unenhanced digitally scanned image of ocean floor in the "famous" region of the mid-Atlantic range.
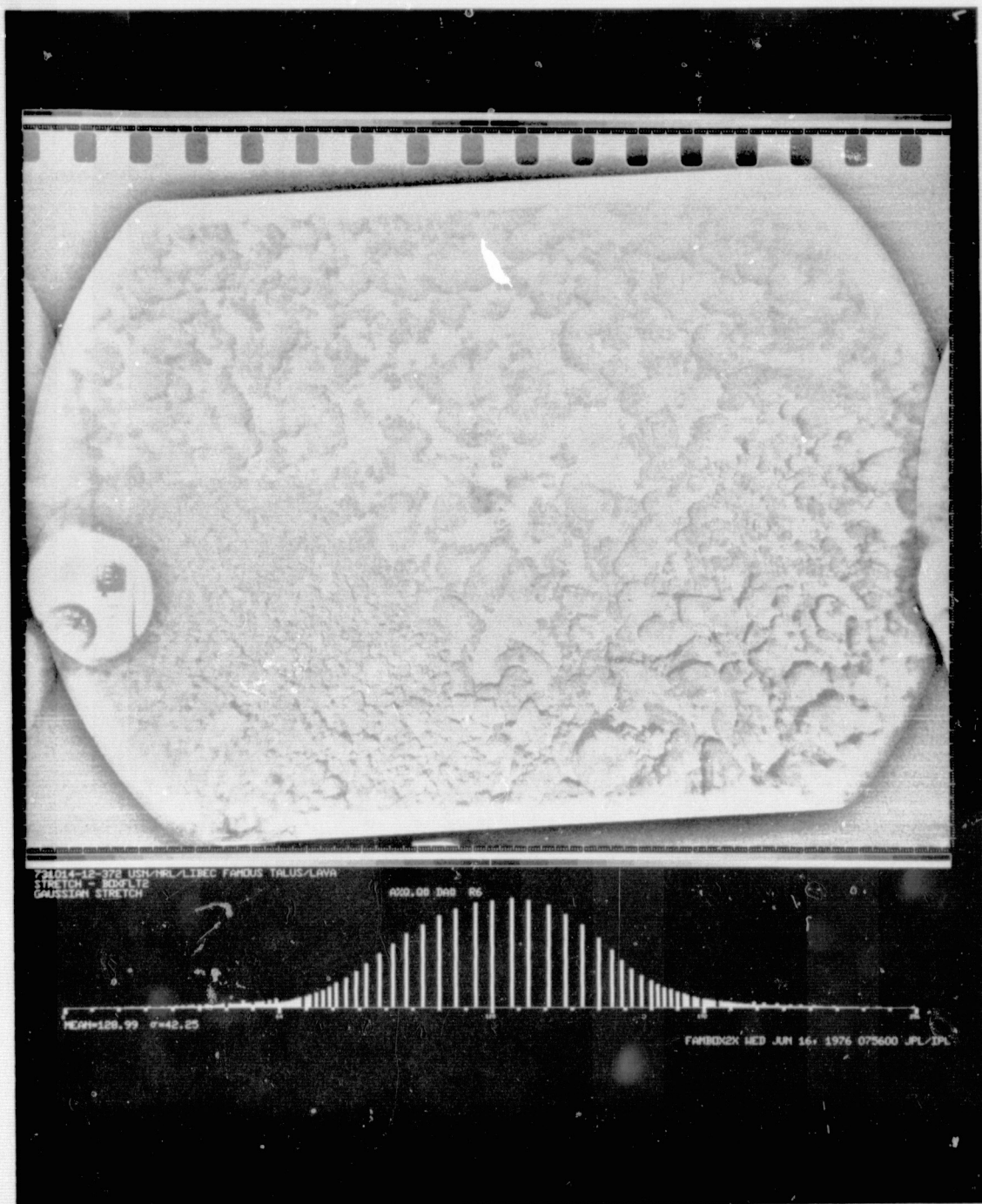
Figure 6b. The same image after Gaussian contrast enhancement.
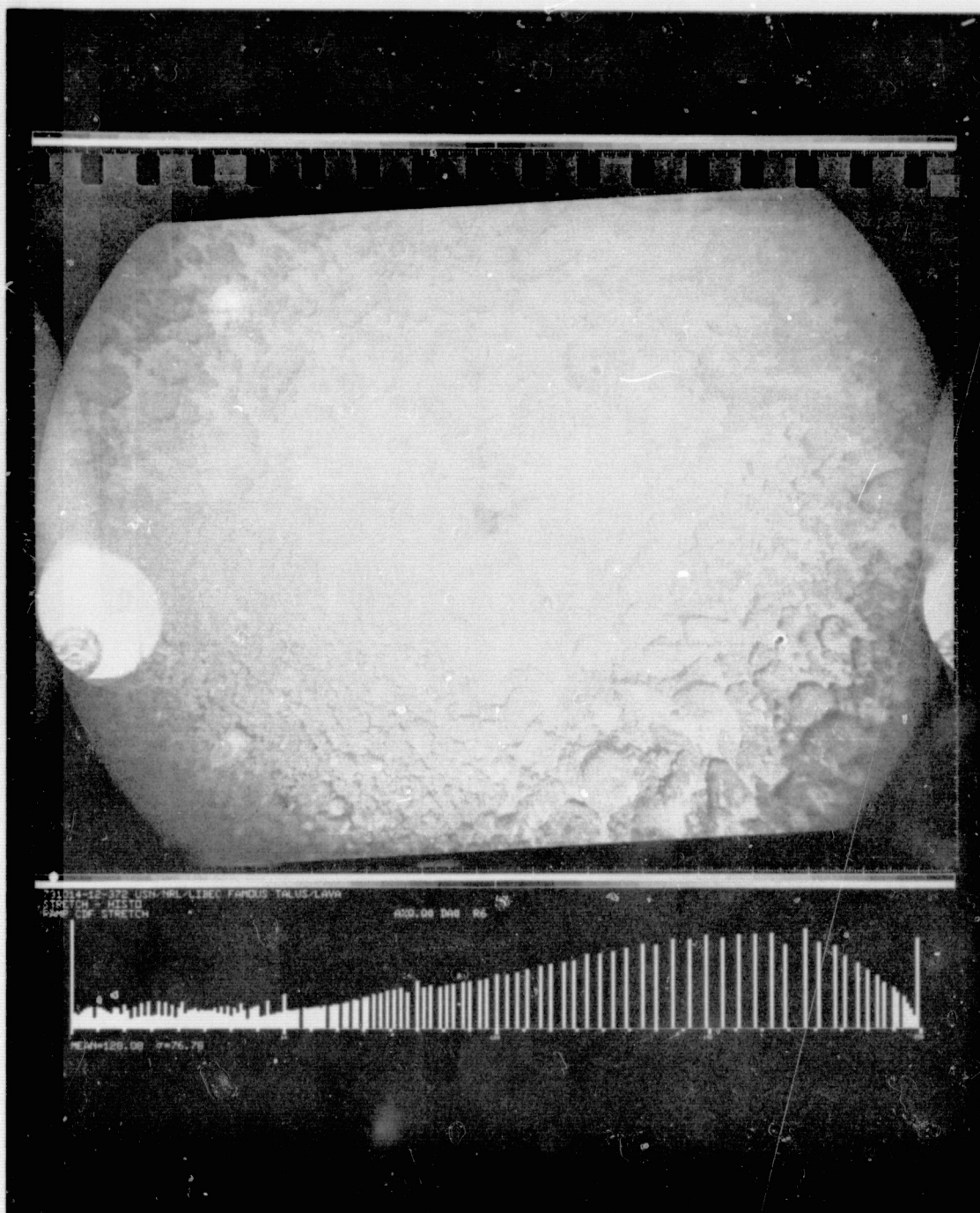
Figure 6c. The same image as Fig. 5a after uniform output distribution contrast enhancement.
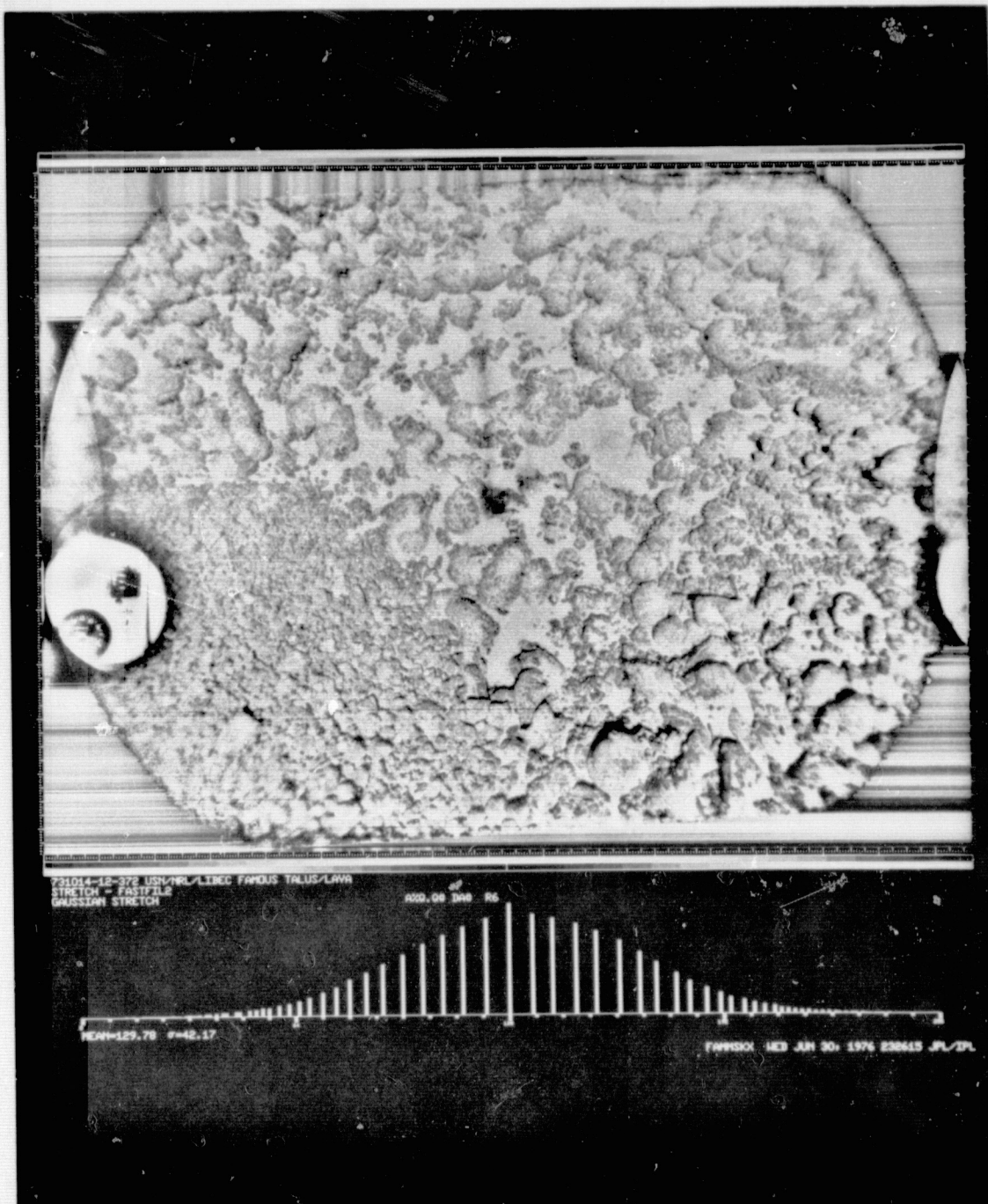
Figure 6d. High pass filtered version with Gaussian stretch applied.

ELEMENTS OF AN IMAGE-BASED INFORMATION SYSTEM

Albert L. Zobrist, Ph.D.

Jet Propulsion Laboratory
Pasadena, California

ABSTRACT

There is a pressing need for systems which can manage spatially-referenced data and perform certain types of spatially-oriented processing. As an example, census data is spatially referenced, and there is a standard repertoire of computational steps performed by users of that data such as aggregation of variables over regions defined by polygons. The IBIS (Image Based Information System) system extends the capability of present systems by adding a new datatype, the image raster, in such a way that it can be used with tabular data bases. Many of the existing computational procedures are simplified or made more accurate when converted to a digital image processing mode of operation and several new computational procedures are made possible. In addition, new methods of data capture and new sources of data are added (e. g., photo-scanning and satellite imagery).

Several technical problems have been overcome to achieve a workable system. First, digital image file handling, image manipulation and image processing capabilities must be provided. Second, image data must be registered or indexed to spatially-referenced tabular data so that processing steps which involve both types can operate. Third, a data interface must be provided between the different datatypes so that the results of processing can be represented. Finally, image processing analogs must be developed for existing geo-base file computational steps (e.g., overlay, aggregation, crosstabulation, etc.).

A user request is given to the IBIS system by means of a language which is translated into the host machine job control language. The object code can then invoke system functions or processing modules. This organization makes the system flexible and easily extendable.

The system is now in use on a test basis. Its first application involves the tabulation of land use by census tract for Los Angeles and Orange Counties, California, from information contained in satellite imagery. Test cases for more advanced problems have been run and spatial modelling capabilities are being developed.

I. INTRODUCTION

The purpose of this report is to communicate the functional requirements and system description for the Image Based Information System (IBIS) developed at the Image Processing Laboratory at the Jet Propulsion Laboratory. The initial motivation for the development of IBIS was to introduce remotely sensed imagery into the mainstream of data processing application. A basic example is that of processing a LANDSAT thematic map showing land use or land cover in conjunction with a census tract polygon file from the Census Bureau Urban Atlas to produce a tabulation of land use acreages per census tract. In the analysis of the steps necessary to achieve this basic capability, two facts became clear. First, a large number of image processing and data manipulation capabilities would be needed for even the simplest case. Second, with the proper design, the minimal system can be extended into a general information system with novel features and capabilities. The term Image-Based Information System has been adopted because the image datatype and image processing operations are crucial to many of the new capabilities. Some elementary applications of IBIS are described in a separate report in this conference. [1] Here, emphasis is placed upon giving a general description of IBIS which is independent of any particular application. Applications will be mentioned for illustrative purposes only.

II. FUNCTIONAL REQUIREMENTS

1. Data Management

There are a number of ways to encode spatial data. [2] It is possible, however, for illustration purposes, to dichotomize referencing systems into nominal and ordinal, and data types into tabular, graphical, and image. Ordinal systems reference data by the actual geographical coordinate values. Thus, natural resources such as forests, rivers, crop land, and geologic formations are mapped with selected identifiers (total area, boundary, centroid) referenced to latitude and longitude or other selected geographic coordinate system. Nominal systems are "name referencing", i. e., data or information is referenced to a name-designating system. Any district-based referencing convention, such as census tract, sewer district, township, or transportation zone, assumes the operator knows where each administrative area is located and leaves the analysis of contiguity-effects, etc., up to the individual. [3] Problems involved in conversions between ordinally and nominally referenced data and have, with the exception of the Census DIME (Dual Independent Map Encoding) File system, forced the conversion of all data to one reference system or the other. [4]

Of particular interest here is the use of image data type. Until recently, the image format has been used primarily as a computer processable equivalent of a photo, with the value stored in each cell of the image representing a shade of grey or a color. But if the image is of a geographical area, then the value in the cell can be a datum for the area corresponding to that cell. A principle advantage of the image representation is that data for a geographical point can be accessed immediately by position in the image matrix. Figure 1 illustrates the calculation of memory address of the data value from a latitude-longitude pair. The image datatype seems to be

LOCATION FORMULAS:

$$X = \left[A \times LAT + B \times LONG + C \right]_{\text{NEAREST INTEGER}}$$

$$Y = \left[D \times LAT + E \times LONG + F \right]_{\text{NEAREST INTEGER}}$$

MEMORY ADDRESS = BASE + KX + Y

Fig. 1. Image Matrix as a Data Representation

a powerful and general representation for spatially distributed data, and the range of uses can be divided into several broad categories.

- **Physical Analog.** The pixel value represents a physical variable such as elevation, rainfall, smog density, etc.
- **District Identification.** The pixel value is a numerical identifier for the district which includes that pixel area.
- **Class Identification.** The pixel value is a numerical identifier for the land use or land cover, or for any other area classification scheme.
- **Tabular Pointer.** The pixel value is a record pointer to a tabular record which applies to the pixel geographical area.

This range of uses requires that the system be able to handle images composed of words of varying length. For example, to identify census tracts in Los Angeles County requires 1500 different pixel values and elevation maps can require 15000. This is more than the usual 256 grey levels handled by photographic image processing systems.

For each data type, the overriding consideration is usually getting the data into the system, for it is in this area where greatest costs and difficulties are usually incurred. Figure 2 depicts data input as a three stage process. The first stage, called data capture, includes all operations up to the point where a data file is computer readable. Data capture costs are enormous for many basic kinds of data, for example, demographic and economic data gathered by the U.S. Bureau of the Census. These data are then made available on computer tape at nominal cost to any user. Another common method of data capture is coordinate digitization of boundaries or linear features from a map. The map is not computer compatible but the digitizer output is. Manual photointerpretation (for example, manual determination of land use boundaries) is a common step prior to coordinate digitizing. LANDSAT

imagery is a particularly attractive data source because it is already computer compatible and gives up-to-date coverage of large areas at nominal cost. Editing costs vary widely depending upon the nature of the data source. Reformatting can also be a major operation where large data files are involved. A good example is community analysis where data are gathered by various districtings (police, fire, sewer, census tract, etc.) but must be reformatted to a common districting so that analysis can be performed (for example, to obtain police calls per capita). The final stage of data input is to obtain a temporal baseline. For example, police calls in 1976 cannot be divided by 1970 population to obtain a per capita figure because of a temporal difference. Establishing a temporal baseline involves projection and modelling of the initial data to one or more common points of time.

2. Data Analysis

Once a working data base is set up, provisions must be made for information retrieval, information analysis, and report generation. Operations here are usually of a much smaller scale than data capture in terms of time and cost, but there is instead a question of flexibility and ease of use, and a question of system cost. All data analyses can be laid out as a sequence of primitive steps, thus, a functional requirement is that an adequate set of primitive operations can be implemented. Mathematical and statistical analyses of tabular files are well-understood, and packaged systems can easily be interfaced. The open question is whether geo-based file computational steps can be implemented. Some examples of these are: 1) Given a point and a district, does the point lie within the district. 2) Given a point and a district file, which district contains the point. 3) Given a particular district in a district file, what are its neighbors. 4) Given a district file and an area classification file, what are the acreages of each area classification in each district. 5) Given two district files, one major and one minor, what
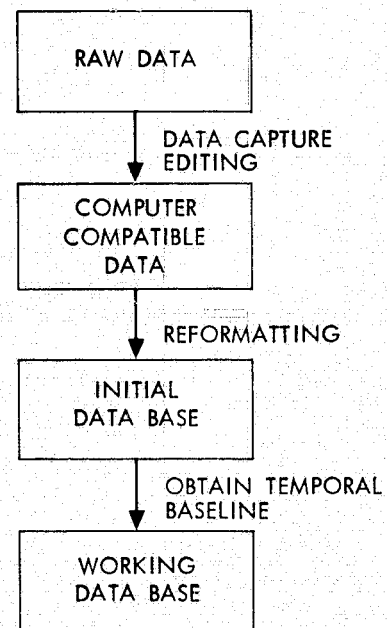

Fig. 2. Data Input Stages

are the proportions of each minor district in each major district. 6) Given a district file and a line segment, what are the mileages of the line segment in each district. 7) Given a point p and a point file, what is the point in the point file nearest to p. 8) Given a point p and a point file, what is the distance from p to the nearest point in the point file. 9) Given a point and a line segment file, which line segment passes closest to the point. 10) Given a density map and a district file, what are the volumes in each district (spatial integral of density). 11) Given a district, what is the centermost point (an inside point which is farthest from the boundary). And for the connoisseur: 12) Given a district file, assign four colors to the districts so that a map can be produced with adjacent regions always a different color.

The preceding list is just a sample of the sorts of spatial or geometric calculations which need to be performed by a comprehensive geo-base information system. More complex operations will usually be implemented as a sequence of these primitive operations, but because of the magnitude of the data files and because of iteration due to modelling, compute time can be a serious problem. A method which solves one of the primitive problems in 0.1 second may seem usable, but not if it has to be performed ten million times for a particular application.

It is worth noting here that many of these operations are difficult and time consuming if the working data base is in polygon or graphical format (i.e., lines are specified by their end points and a district is given by a sequence of line segments). In particular, the operation called polygon overlay which solves primitive problems four and five is extremely difficult to perform on large files in graphical format. If the files are in image format, then polygon overlay becomes a simple counting operation (Fig. 3).

A different area of concern is the interfacing of nominal and ordinal geo-referenced files. Given that part of the working data base has been

obtained (or reformatted) into an ordinal format, the major part of the data base will probably still be in tabular form. As an example, an air pollution density image might be interfaced with a population table to obtain a measure of health effect. Speaking more generally, the system must be able to perform operations on mixed data types, thus allowing the working data base to be built from raw data in its most natural form.

An operation of special importance is cross-tabulation, mentioned previously in connection with data management. It converts data aggregated by one district convention to an aggregation by another district convention. The operation itself is trivial since it involves multiplying by a set of factors which measure the percentage subareas of one districting in the other. The cross-tabulation factors are derived by polygon overlay and may also be modified by a density estimate of the variable being crosstabulated. A comprehensive system should be able to represent district data sets in such a way that they can easily be updated and the factors rederived. This means that the tabular data can be kept in its natural unit of aggregation in the working data base and quickly crosstabulated to whatever districting is needed as analysis proceeds.

Because the image datatype is used, capabilities for digital image file handling, image manipulation, and image processing are required. Thus, the IBIS system has been built upon an existing image processing system, VICAR (Video Image Communication and Retrieval), developed at JPL.[5] Certain basic image processing operations are absolutely essential. One must accomplish image-to-image registration, whereby images of different scale, rotation, or map projection are superimposed precisely enough so that corresponding pixels represent the same geographic location. Rubber-sheet registration is almost always necessary to achieve the needed degree of accuracy. On the other hand, it is anticipated that even esoteric image processing operations, such as convolution smoothing will be useful for certain types of applications. The conclusion here is that any image-based information system must contain a powerful image processing subsystem.

Finally, the system must be capable of computations which will allow it to run a variety of models. The key here is the introduction of the image raster as a spatial data representation. Contiguity effects can then be handled because adjacent areas are contiguous in the image raster (in terms of their computer address) whereas, with other file representations, adjacent areas are accessed by file searching. Basic operations, such as spatial integration with reporting by district are greatly simplified with an image representation. More esoteric models, say, involving cellular transformations[6] or diffusion processes may already use a matrix on raster representation but will be helped by the synergism of placing their specialized capabilities into a general and comprehesive data management system.

III. SYSTEM DESCRIPTION

The hardware configuration presently used by IBIS is diagrammed in Figure 4. The operating system is OS 360/MFT and all IBIS routines are designed to run in a 150 kilobyte region. Present

GRAPHICAL FORMAT:

| DISTRICT FILE 1 | DISTRICT FILE 2 | MUST CALCULATE AREA OF M x N INTERSECTIONS OF POLYGONS |

(M DISTRICTS) (N DISTRICTS)

IMAGE FORMAT:

DISTRICT IMAGE 1
DISTRICT IMAGE 2

MUST COUNT K x L PIXEL PAIRS
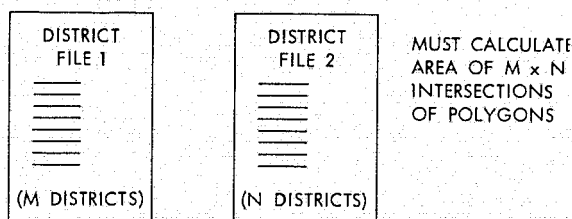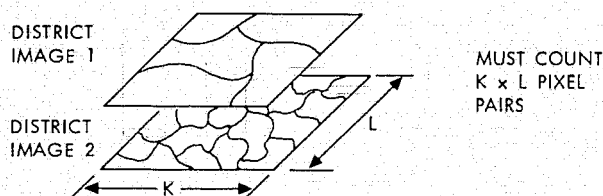
Fig. 3. Comparison of Polygon Overlay Techniques
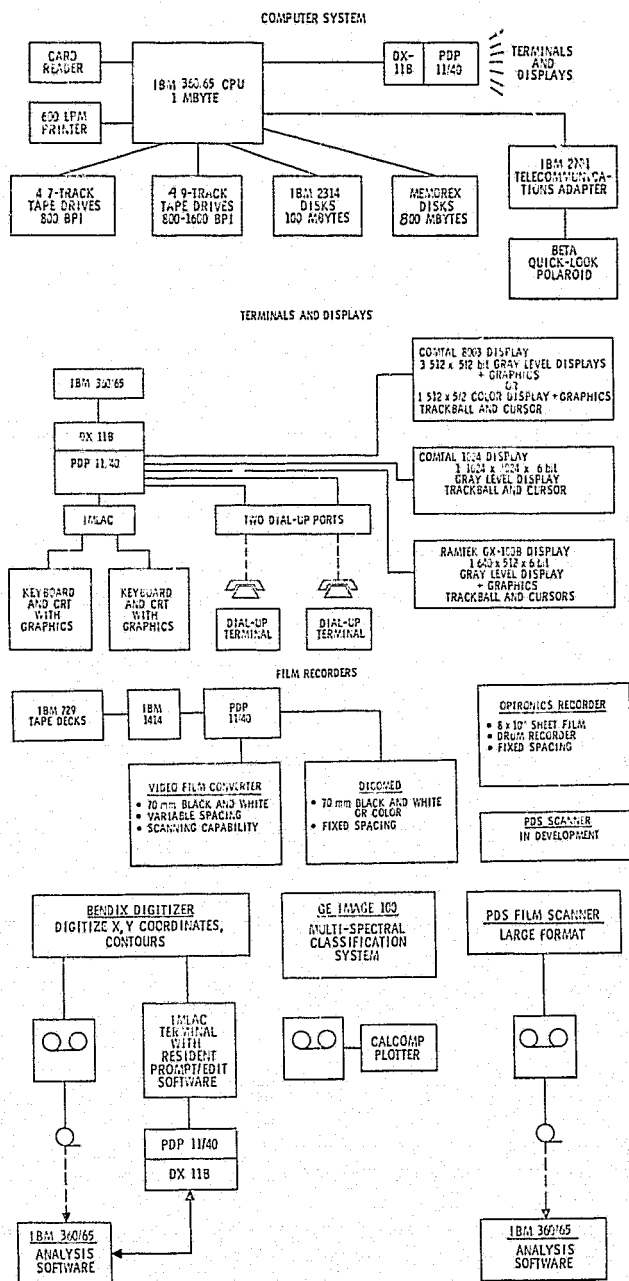
Fig. 4. Hardware Configuration

applications use less than one twentieth of the total machine time, so this hardware configuration would be suitable for a large scale data processing operation at the national level. A minicomputer version is partially complete. Although the total cost of an IBIS system is high, the unit cost of operation is low, especially when compared to existing methods of performing the types of tasks the system is intended for.

The user request is given to the IBIS system by means of a language which is translated into the host machine job control language. The translated code can then invoke system functions or processing modules. This organization makes the system flexible and easily extendable. The system software consists of a number of FORTRAN modules and a relatively small nucleus. All of the factors mentioned here (modular design, use

of FORTRAN, good interfaces to user and hardware) make technology transfer more feasible.

The processing modules constitute a primitive set of functions operating on the various datatypes to achieve the functional requirements set out in the previous section. Figure 5 is a schematic layout of the central part of the IBIS system, emphasizing the interfaces between image, tabular and graphical datatypes.

The first set of routines convert graphical (polygon) data files to digital image descriptions of regions or areas.

VTRACTI, SCRBGEN, VSCRBGEN - All three programs basically have the same function, convert polygon files described by x, y coordinates to a VICAR record format, but each uses a different input type:

VTRACTI	-	Census Tracts
SCRBGEN	-	Cards
VSCRBGEN	-	Tape

POLYREG - Performs rigid rotation and scaling of polygons prior to input into POLYSCRB.

POLYGEOM - Performs rubber sheet mapping on polygons prior to input into POLYSCRB.

POLYSCRB - Scribes lines into an image file, by setting the boundaries of polygons to a particular DN (grayscale) value. The routine was especially designed for the conversion of thousands of lines and has parameters for chaining lines and closure of polygons.
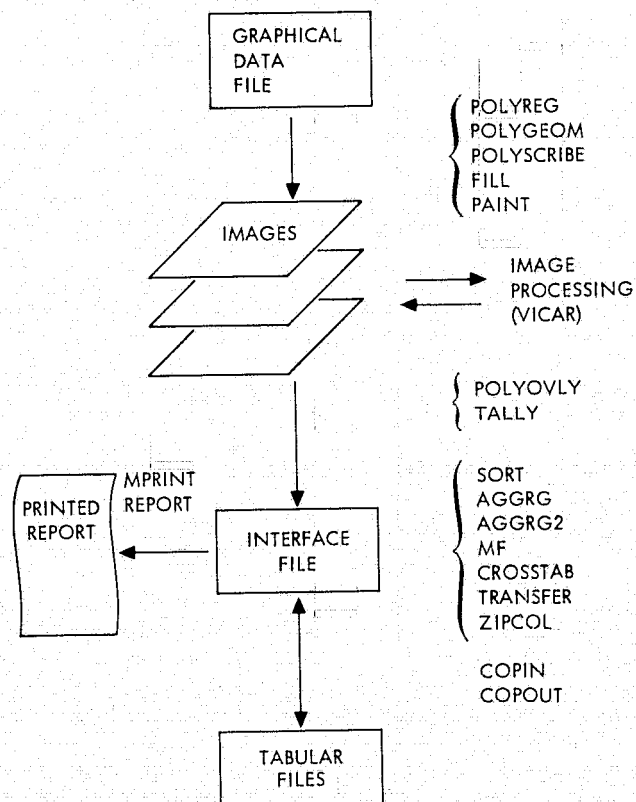
Fig. 5. IBIS Schematic

FILL - Fills holes, thickens lines and/or removes noise with variable thresholds and window sizes for enhancement of scribed polygons prior to PAINT.

PAINT - Converts an image with scribed polygons into a multi-color map where the DN corresponds to a map color. The routine will handle up to 30,000 polygons or regions.

Moving from top to bottom in Figure 5, the VICAR image processing routines are next. The VICAR system presently has over 100 routines for the manipulation of images of varying word length. Of special interest to modellers is the following routine.

F2 - Performs array arithmetic on a pair of images in byte or halfword format. The function to be applied is specified by a FORTRAN-like expression.

Polygon information extraction routines perform the overlay operation on two sets of polygons, the reference polygon created in paint and digital image data. Each pixel on the image file is matched with the corresponding pixel on the reference file and stored by pixel pairs. The following programs are used in the overlay process.

POLYOVLY - Produces a histogram (or pixel count) or registered images by DN values for single pixels or a joint histogram by DN pairs as generated in the overlay procedure. Since the problem of storage is critical to this process, the user has three optional storage methods specified in the parameters.

TALLY - Produces a histogram similar to POLYOVLY except that the DN values in one of the two images are summed.

The results of POLYOVLY and TALLY generate data files stored by column. The column organized file is called the interface file. This file serves as an interface between image and tabular data sets. The interface file is manipulated to produce a report by the following routines:

SORT - Sorts the interface file into ascending or descending order according to one or more columns.

AGGRG - Aggregates columns of numbers using a designated column as an index for summation in other columns.

AGGRG2 - Summarizes and collates columns of numbers to produce a single row for each control value in the column which is used as an index.

MF - Performs column arithmetic given an arithmetic expression in the parameter field which denotes columns such as:

"C5-C2/C3+100+SQRT(C2)".

CROSSTAB - Tabulates information referenced to one polygon districting to another polygon districting through the polygon overlay technique.

TRANSFER - Change vertically aligned columns of data as produced in POLYOVLY and used in previous routines to smaller vertical columns based upon data values (e.g., land use).

ZIPCOL - Substitutes column index values with user district names or numbers.

After processing by the above routines, the interface file is ready for report generation.

MPRINT - Prints out the interface file according to a relatively simple format.

REPORT - Prints out the interface file with user specified titles, column titles, paging, spacing, etc. Subtotaling on a given control column can be requested. There are numerous features for formatting and for printing alphabetic data.

To interface tabular files and also to give tape output as an alternative to report generation, the following routines are used.

COPIN - Copies columns of a tape file into columns of the interface file according to an index column in the interface file. The index must be present in the tape file. The files must be sorted according to the index and a merge is performed.

COPOUT - The reverse of COPIN. Columns are written from the interface file to the tape.

BIBLIOGRAPHY

1. Bryant, Nevin A., "Integration of Socioeconomic Data and Remotely Sensed Imagery for Land Use Applications," Proceedings Caltech Conference on Image Processing, etc. October 1976.

2. Deuker, K. J., "A Framework for Encoding Spatial Data," Geographical Analysis, Vol. 4 (1972), pp. 98-105: Deuker, K. J. "Urban Geocoding," Annals, Association of American Geographers, Vol. 64 (1974), pp. 318-325; Werner, P. A., "National Geocoding," Annals, Association of American Geographers, Vol. 64 (1974), pp. 310-317.

3. Cooke, D., Storage and Retrieval of Geographic Data, URISA, SIG/GBF Technical Session, Atlantic City, NY, 1973.

4. U.S. Bureau of the Census, Census Use Study: The DIME Geocoding System, Report No. 4, Washington, DC, 1970.

5. Billingsley, F. C., "Review of Digital Image Processing," Proceedings of EUROCOMP '75, London, England, September 1975.

6. A term coined by Prof. Waldo Tolber, University of Michigan.

COMPUTER TECHNIQUES FOR GEOLOGICAL APPLICATIONS

James M. Soha
Alan R. Gillespie
Michael J. Abrams
Daryl P. Madura

Jet Propulsion Laboratory
Pasadena, California

ABSTRACT

Remote sensing imagery has proven to be an important new data source for investigations in geology. Digital image processing is in turn a powerful tool for analyzing and extracting information from this image data. This paper describes several of the basic image processing techniques now being used in geologic work at JPL. Particular emphasis is placed on the processing and display of Landsat multispectral data.

INTRODUCTION

The recent widespread availability of remote sensing data has opened new horizons for investigators in the earth sciences. Pictures such as those returned by NASA's Landsat I and II satellites are beneficial to geologists in two principal ways. First, the overall, synoptic view provided by these images presents large as well as small scale structural elements within their contextual and relational framework. A second important attribute of these images is the multispectral information they contain, allowing differentiation of surface materials according to variations in spectral reflectivity.

The quantity of data returned by Landsat is enormous: each Landsat scene contains over six million picture elements in each of four spectrally different, spatially registered images. The fact that this data is digitally encoded makes computer handling and manipulation of the data particularly convenient. Sheer volume of data is not the only reason for employing digital image processing in the analysis of Landsat data. Much of the geologically interesting information in a scene is subtle in nature and is simply not visible in the photographic products which can be ordered from the EROS Data Center of the USGS. Careful computer processing and controlled display of the data is necessary to extract this information. As in any other science, geologists want to press the diagnostic and informational capability of their data to the practical limit. For remote sensing, this objective makes digital image processing a virtual necessity.

This paper describes several of the computer techniques that are commonly used at the Image Processing Laboratory (IPL) of the Jet Propulsion Laboratory in support of research in geology involving Landsat and other remote sensing data. First, several of the basic operations with single (black and white) images are discussed. Preliminary cosmetic processing to remote distracting noise and artifacts is an important aid to further analysis. Geometric transformation of images is useful in producing a standardized cartographic display format for comparison with existing maps, or to register multiple images for combined analysis. Careful contrast enhancement is vital for successful display. Filtering is particularly helpful in enhancing structural information. Next, extension of these techniques to multispectral image analysis is described. Finally, the problem of information extraction from multispectral images is discussed. Little attempt is made to present the actual application of these techniques to current geologic investigations at JPL; this subject is addressed by two other papers in these proceedings (Refs. 1 and 2).

COSMETIC PROCESSING

The Landsat MSS (multi-spectral scanner) is composed of six sets of four different sensors, so that six scan lines are acquired simultaneously in each spectral band. Because the six physically different sensors in any given band possess slightly differing response characteristics, it is not surprising that residual striping with a six line pattern remains in the images even after the radiometric correction performed by Goddard. This stripe noise can be distracting, so that some effort to reduce it is worthwhile. While it is not feasible to model it precisely and then remove it, practical procedures do exist for reducing the impact of striping.

The most practical of these approaches is to accumulate separate frequency of occurrence histograms of brightness for each of the six different sensors (these histograms count the number of picture elements having each possible brightness value). Then a different intensity transformation is applied to the sets of lines acquired by each of the sensors. The basic idea is that on the average over an image, each sensor should observe about the same distribution of scene brightnesses. The simplest approach is to force the mean and standard deviation of each sensor distribution to match the mean and standard deviation of the image as a whole. An additional refinement of this method consists of forcing the envelope of each sensor histogram to match the shape of the envelope of the entire image histogram. Figure 1 shows this technique applied to a segment of a Landsat scene including Goldfield, Nevada.

Filtering can also be used to reduce Landsat striping, but is generally less suitable (filtering techniques are discussed in more detail later). One line horizontal "box" filtering will certainly remove striping, but it also removes other desirable vertical frequencies. Vertical convolutional notch filtering is another possibility (primary noise frequencies are 0.166, 0.33 and 0.5 cycles per sample), but a large weight matrix is required to achieve sharp bandpass response. Ringing off sharp brightness boundaries (e.g., clouds) can also be a problem with each of these two methods. Either one or two dimensional Fourier domain notch filtering is also a possibility, although computation time can be a problem.

Striping can also exist in other scanner images (particularly aircraft scanners) acquired with a single sensor (per spectral band) and hence not exhibiting "coherent" six line character. In many cases this striping can also be reduced by using frame average behavior. One technique operates as follows. First, the average brightness value of each line in the image is computed, and the results stored as a new one-column average image. Subtractive box filtering is then used to determine the difference between each element in this one column image and the average of many (perhaps several hundred) values surrounding it. The resulting values represent the variation in average behavior between a given line and those lines surrounding it and can be thought of as additive approximations to the stripe noise. These values are then subtracted from each picture element of the corresponding scan line to reduce the noise. Where the striping is too erratic, sometimes a nearby spectral band can be used to define the average brightness trend. Figure 2 illustrates an application of this technique.

GEOMETRIC TRANSFORMATION

There are several reasons for performing geometric transformations of images. One common reason is to remove relative distortions in a scene, caused by sensor characteristics or scene geometry, which can produce a confusing non-uniformity in an image. One example of this problem (called panorama distortion) is the variation in both pixel spacing and footprint size across a scan line in a scanner image caused by uniform angular sampling and variation in slant range. Sometimes it is desirable to transform an image to a standard cartographic mapping projection to permit comparison with existing maps or to produce a display more meaningful to other individuals. Or, it may be necessary to register two or more images acquired at different times or by different sensors to permit common analysis.

The computer technique used at IPL to perform geometric transformations first requires definition of the transformation in terms of its effect on each pixel. One method of defining a geometric transformation is by a function which specifies displacement and can be evaluated at each picture element. This approach is used for dealing with panorama distortion (a method for removing panorama distortion is described in detail in Ref. 3). Another procedure is to define the transformation by providing specific displacements for a subset of the pixels in the image (usually some sort of regular grid). Displacements for all other pixels can then be obtained by interpolation.

In either case, since the output (transformed) image must consist of a regular square pixel grid whose elements in most cases do not correspond exactly to any pixel in the original image, resampling must be performed. Resampling is the process of selecting a plausible brightness value corresponding to a non-integral pixel location by examining the brightness at nearby elements. Several resampling techniques are available (Refs. 4 and 5). When the original scene is sufficiently bandlimited (containing no power at frequencies greater than 0.5 cycles per sample in the sampled image), perfect resampling is theoretically possible using sinc functions. Even when the sinc

function is truncated at reasonable limits, however, this approach requires far too much computation time to be realistic in any but special circumstances. Practical alternatives in common use include the nearest neighbor algorithm (the brightness value of the nearest pixel is used), bilinear interpolation, and cubic convolution (using the nearest 16 neighbors). At IPL, bilinear interpolation is most frequently used, although a cubic spline interpolating function is sometimes employed in one-dimensional cases (where only four input samples are required to compute each output value).

Removal of geometric distortion from Landsat images has been described elsewhere (e.g., Refs. 4 and 6). The distortions which are generally removed at IPL are earth rotation (introducing a skew into the image), panorama, roll, aspect ratio, mirror scan velocity non-uniformity and "synthetic pixels" (the latter two can be handled using information in Ref. 7). In all cases, the distortions are modelled, a mapping is defined, and resampling is performed. All of these corrections are particularly convenient to perform since they can all be handled with one-dimensional resampling. In fact, at IPL, these corrections are performed concurrent with initial reformatting of the data, before the destriping operations described previously.

The software used at IPL to transform planetary image data acquired with perspective camera systems to cartographic mapping projections (Ref. 8) is not currently applicable to scanner images taken by Landsat, NOAA, or aircraft scanners. At present, visual inspection of these images is used to determine points in common with existing topographic maps. Relative displacements are obtained for a grid of points sufficient to define the transformation. When a sufficiently regular grid cannot be found readily, polynomials are fit to the displacements of the points which are found, and these polynomials are then evaluated at a set of grid points. Figure 3 shows an aircraft scanner image of Goldfield, Nevada, before and after transformation to a Universal Transverse Mercator projection.

Occasionally it is necessary to register two or more images taken at different times. For example, to produce a thermal inertia image (Ref. 2), it is necessary to register a daytime and a nighttime thermal image to obtain a difference picture. In such cases, correlation between the two images can be computed for local areas surrounding hand selected tie points. At IPL, selection of tie points and the computer correlation are performed interactively, and visual feedback is used to assess the success of the correlation. Once a set of displacements is obtained, a geometric transformation is performed on one of the images, using the techniques described previously.

CONTRAST ENHANCEMENT

By far the most commonly applied image enhancement procedure is contrast modification. Sensors typically must be designed to handle a wide range of brightnesses without saturating, so that individual scenes often do not utilize the full available range of pixel DN values. Or, a user may be primarily interested in only portions of a scene and wish to emphasize those features. In either case, it is important to fully utilize the available

dynamic brightness range in the display medium. While this contrast adjustment could be performed by varying the transfer function of the display device, it is generally preferable to leave the parameters of the playback device fixed and alter the contrast of the image while it is in digital form.

Contrast enhancement consists of an intensity transformation applied to the image which maps brightness values in the input picture to other values in the output picture, usually according to some formula. A useful analytic device for understanding the effects of contrast enhancement is the intensity frequency histogram. Figure 4a illustrates the histogram of a section of a Landsat frame of a portion of the Chile-Bolivia border with no contrast enhancement. Clearly, only a fraction of the available range is utilized. If this digital image is converted to a film transparency using a fixed transfer function to produce a product with a density range suitable for color reconstruction (e.g., 0 DN = 1.3, 255 DN = 0.15), then a picture similar to that of Fig. 5a results, which is clearly unsatisfactory.

Several contrast enhancement techniques (i.e., methods of specifying the intensity transformation) are in common use. The simplest of these is a linear contrast enhancement in which some value L is mapped to 0 DN, some value H is mapped to the maximum output value (usually 255), and intermediate values are scaled proportionately. Input values extreme of L and H are saturated to black and white respectively. The two stretch limits L and H can be chosen after inspecting the input histogram. Alternatively, a program can accumulate and scan a histogram to determine the two limits defining the stretch. Often this is done by specifying what percentages of the picture are to be saturated black and white. The program then integrates inward from the extremes of the histogram until those percentages have been reached to determine L and H. Another approach is to assign the two input values specified by percentages of the histogram to output values other than extreme black or white. Figure 5b illustrates the results of an automatic linear contrast enhancement in which 4 percent of the histogram was saturated to 0 and 255. The resulting histogram is shown in Fig. 4b.

Non-linear intensity transformation can also be used. Piecewise linear transformations (often called "table stretches"), in which the transfer function is specified for certain values and interpolated linearly between these values, is one convenient method. Carrying this approach to its logical extreme yields the most general form of contrast stretch in which a particular output value is assigned for each possible input value. Contrast modification programs can be designed to produce these tabular transfer functions according to some user specified fundtion, such as logarithmic or power law. Alternatively, after scanning the input histogram, a program can specify a tabular transformation which will produce an output histogram of a certain character, such as a uniform or a Gaussian distribution. Forcing a uniform distribution in the output picture results in the greatest contrast enhancement being applied to the most populated range of brightnesses in the input image. This property makes the uniform distribution stretch particularly useful as a quick look method for evaluating the results of a previous

processing step. The principal difficulties with this approach are that it is sometimes too harsh and that it can result in severe compression of the histogram and a resulting loss of brightness in less populated areas of the input distribution (such as the bright and dark "tails") which are sometimes of particular interest. Figure 5c shows the result of a uniform distribution stretch applied to the Chile-Bolivia test frame. Notice that structural detail of the volcano near the center has been virtually saturated dark. Definition of detail within the extremes of the histogram tends to be preserved or increased more effectively with a contrast enhancement which forces the output histogram to be a Gaussian of a user specified width (standard deviation). There will be correspondingly less emphasis in the central (usually most populated) brightness zone. To the extent that the input distribution is normal, a Gaussian contrast enhancement resembles an automatic percentage saturation linear stretch. A Gaussian stretch can be particularly useful in dealing with a biased, non-symmetric input histogram, such as a log normal distribution. Figure 5d illustrates a Gaussian stretch applied to the test image.

The most effective contrast enhancement in any particular case depends of course on the character of the input histogram and which features in the scene are of greatest interest to the user. For this reason, best results are usually obtained by an experienced human analyst inspecting the image histogram and devising a stretch to meet his needs. The automatic contrast enhancement algorithms are of greatest value when time or sheer volume of work precludes careful human analysis of each input histogram.

FILTERING

Filtering can be thought of as any process which differentially modifies image content, tending to emphasize desirable features while suppressing less desirable ones. Most often, filtering is visualized in terms of spatial frequencies, as expressed by the Fourier transform. In these terms, filtering is accomplished by modifying the modulus and/or phase of some or all frequency components.

Linear filters can be implemented either by multiplication in the frequency domain or by the mathematically equivalent procedure of convolution in the spatial domain (Ref. 9). Pratt calculates (Ref. 10) that whenever the convolution weight matrix exceeds about 10 pixels square, frequency domain filtering, including the necessary Fourier transforms, is faster, assuming the transform program is efficiently coded. At the IPL, the existence of a special high speed convolution hardware unit raises this threshold, so that convolution filtering is utilized much more often. Also, Fourier transforms of large images such as Landsat (which also lack a convenient power of two in their dimension), pose logistical problems under the IPL software system.

Notch filtering, i.e., the reduction or removal of specific frequency components, can be helpful in removing coherent noise from an image. Such noise can be introduced by electronic defects in the sensor. Figure 6a shows an image acquired by an aircraft scanner in which coherent noise is a problem. A one-dimensional vertical power

spectrum (Fig. 6b) illustrates the noise spikes. One dimensional vertical notch filtering was then applied to remove these noise frequencies, with the result shown in Fig. 6c (actually for ease in processing, the image was rotated 90 degrees and horizontal filtering performed).

Filtering can also be used to boost high frequency content, thus improving definition of edges in an image. This is particularly useful when analysis is aimed at obtaining structural information from the picture. Sensors are generally less responsive at higher spatial frequencies. For Landsat I, this fall off in response has been measured (Ref. 4) by differentiating across a sharp boundary and obtaining the Fourier transform of this measured point spread function. Response was found to be about 30 percent at the highest spatial frequencies. Where this frequency response is precisely known, e.g., via a Modulation Transfer Function (MTF), restoration filtering can be performed. For a noise free picture, this would be done by multiplying the image transform by the reciprocal of the MTF. In practice, there is noise present, so Weiner filtering, designed to minimize mean square error (Ref. 11), is generally used, so that

$$R(\omega_x,\omega_y) = \frac{M(\omega_x,\omega_y)\ I(\omega_x,\omega_y)}{M^2(\omega_x,\omega_y) + \left(\frac{1}{S_n^2}\right)} \qquad (1)$$

where R is the reconstructed transform, I is the transform of the input image, M is the MTF, and S_n is a measure of the system signal-to-noise ratio. At IPL, S_n is generally approximated by a constant. A typical value for S_n in Landsat images, measured by principal component analysis (described later) is 8.5. Often precise reconstruction is less important than merely achieving reasonable edge enhancement, and in this circumstance, good judgement is sufficient in designing a high frequency boost filter.

Low frequency (large scale) intensity variations can sometimes tend to obscure the more interesting local detail in an image. When this occurs, low frequency notch filtering can be useful in obtaining a more uniform display of image detail. While Fourier domain notch filtering can be used, a much faster convolution approach is available. This technique, sometimes called subtractive box filtering or simply box filtering, involves computation of the local average brightness in an M x N rectangular area centered at each pixel. This average, or low pass, value is then subtracted from the corresponding input pixel to yield the output high pass value, according to

$$y_{ij} = s(1-f)(x_{ij} - \overline{x}_{ij}) + sfx_{ij} + C \qquad (2)$$

where y_{ij} is the output filtered value, x_{ij} is the is the input, and $\overline{x}_{ij}$ is the local average value corresponding to line i and sample j. A scaling factor s is used to reduce integer roundoff effects, while C is an offset (bias) value which permits display of both positive and negative variations. A fraction f of the original image can be optionally retained in the filtered output. A recursive implementation, involving only about six add/subtracts and a divide for each pixel, can

be used to determine the low pass value (Ref. 12). The frequency characteristics $H(\omega_x,\omega_y)$ of this filter look like

$$H(\omega_x,\omega_y) = 1 - \left(\frac{sinM\omega_x}{M\omega_x}\right)\left(\frac{sinN\omega_y}{N\omega_y}\right) \qquad (3)$$

Large filter sizes (i.e., values of M and N) on the order of 51 to 201 are commonly used, and for these sizes, only the few lowest frequencies are effectively notched out. Figure 7 shows box filtering applied to remove shading from the noise removed image of Fig. 6. A box size of 151 pixels was used. The problem of applying this technique to multispectral images destined for color compositing is described later.

The box filtering technique can also be used for edge enhancement. If a small filter size (e.g., 5x5) is run and then appropriate percentages of the high pass filtered and the original image are combined, a "pseudo-MTF" filtering operation results. Figure 8 illustrates the technique applied to a segment of a Landsat frame of Iceland.

Caution must be exercised in applying the box filtering procedure, and particularly in interpreting the results, since artifacts can be introduced (Ref. 13). Perhaps the most offensive of these is a preferred directional enhancement, which is particularly noticeable in smaller one-dimensional filters. Figure 9 illustrates the problem, which apparently is caused by the side lobes in the (sin ω)/ω frequency character of the filter. When larger, square filters are used, the problem is minimized.

Figure 10, showing the Altyn Tagh fault zone in China, illustrates the improvement which can be obtained when filtering is performed carefully. Here, low frequency notch filtering and high frequency enhancement were combined.

MULTISPECTRAL ENHANCEMENT AND COLOR DISPLAY

The use of color provides a dramatic increase in the amount of information that can be displayed in an image. Color compositing is also a convenient way to display multispectral images (assuming spatial registration). A different spectral band can be displayed in each of the primaries (blue, green and red) in an additive compositing system, or their complements (yellow, magenta and cyan) in a subtractive system. Variations in spectral response of various materials in the scene then show up as color differences in the composite image, even through colors in the image may bear no similarity to the actual colors of these materials.

Functions of several spectral bands can also be displayed as color composites. Ratios of two spectral bands have proven useful in geologic analysis (Ref. 14) largely because intensity variations due to topography are suppressed so that subtle spectral variations are more apparent. Three different ratio images can be composited in the same manner as three raw spectral bands. Ratio color composites are particularly convenient for displaying Landsat data since all four bands can contribute to the display (e.g., blue = 4/5, green = 5/6, red = 6/7).

Producing an effective display in a color composite image involves several problems which parallel those of black and white image enhancement, but are more complex since more than one component is involved. Contrast enhancement is almost always required. Occasionally filtering is needed. And a problem unique to color is the selection of appropriate functions of the original data which will produce the most informative display. Before discussing these problems, it will be helpful to discuss some color coordinate systems and methods of analyzing color.

Color Coordinate Systems

When three images are used to modulate light in primary colors to produce a color additive composite, the brightness values in each of the component images are called tristimulus values. When two different display devices are employed and use a slightly different set of primaries, then somewhat different tristimulus values are required to obtain equivalent colors. Given the nature of the primaries, it is possible to find transformations (Refs. 15 and 10) relating the two sets of tristimulus values. When an image is displayed temporarily on one device, e.g., a television monitor, but a final product is produced on another device, such a transformation can be useful in maintaining visual consistency.

Another useful set of coordinates is the chromaticity coordinate system. If B, G and R are the tristimulus coordinates (i.e., the original component image brightness values), then

$$b = \frac{B}{B+G+R} = \frac{B}{I}$$

$$g = \frac{G}{B+G+R} = \frac{G}{I}$$

$$r = \frac{R}{B+G+R} = \frac{R}{I}$$

$$I = B + G + R$$

(4)

are chromaticity coordinates. Since $b + g + r = 1$, these three values are not independent. Usually g, r, and I are used as an independent set. In this system, perceived color has been relegated to only two dimensions, with intensity (I) being the third.

A chromaticity diagram or distribution is a useful device for assessing image color characteristics, much as a histogram is helpful in analyzing black and white pictures. Figure 11 illustrates the format of a chromaticity diagram. The two color components g and r are plotted in a cartesian system. Points within the triangle represent colors which are realizable with the three primaries. A color composite image can be scanned and the frequency of occurrence of the various (g,r) combinations counted and stored as a brightness value at the corresponding point in the chromaticity diagram. This chromaticity distribution can then be displayed as an image and used to evaluate color content.

Many other color coordinate systems have been defined and used for a variety of applications (again see Refs. 15 and 10). One other system which has proven useful at IPL is a form of hue, saturation and intensity (H,S,I) system (Ref. 4). Perceptually, hue defines the basic color (e.g., blue or yellow), while saturation measures the degree of whiteness (low saturation indicates pastel colors). If the tristimulus coordinates (B,G,R) are viewed as occupying the positive octant of a three-dimensional cartesian space, then hue and saturation are best visualized by imagining a superimposed cylindrical coordinate system with the origin coincident with the tristimulus origin and with a north polar axis that is identical to the achromatic (B = G = R) axis. In this cylindrical system, hue is the longitude of the point in the color space, while saturation is the colatitude (angle away from the achromatic axis subtended at the origin), normalized by the maximum saturation possible at that intensity (this varies since each tristimulus coordinate has a maximum value). Intensity is ideally the Euclidean distance of the point from the origin, but for computational reasons, is defined essentially the same as for chromaticity coordinates [i.e., I = (B + G + R)/3]. An H,S,I system can be convenient to work with since these quantities are all important parameters in human visual perception. Manipulating colors in this space can thus produce results which are more predictable in appearance.

Contrast Enhancement for Color Display

As for black and white, images must be contrast enhanced to produce a good color display. To obtain a good balance and range of colors in the image, it is desirable to center the chromaticity distribution about the achromatic point and attain a reasonable spread or variation in saturation. Best results are also usually obtained when the basic shape of the chromaticity distribution is not distorted too severely. A non-uniform shape in this distribution often results naturally from the prevalence of a certain material or vegetation regime, and distorting the distribution can obscure this natural differentiation. Thus the apparently obvious procedure of forcing a uniform distribution of points on the chromaticity diagram is often not productive. As well as obtaining spread in the color coordinates, contrast should also be increased in the intensity component, usually in a monotonic relationship to the intensity in the raw scene (although an exception will be described).

It is not necessary to transform color images to chromaticity coordinates to obtain successful contrast enhancement. Contrast stretching of the individual tristimulus component images (i.e., blue, green and red) can produce good results, especially when certain helpful rules of thumb are followed. Generally, good results are obtained by stretching the three component images so that the shapes of the three histograms look similar (to produce good color variation) and each individual image has appropriate contrast when viewed as a black and white image (to insure adequate intensity variation). This can be accomplished either by hand selected stretches for each band (once again, manual selection by an experienced analyst usually produces superior results) or by use of one of the histogram normalization stretches. Use of the uniform

distribution stretch tends to produce high saturation (due to the wide dispersion this technique introduces into each component), which can be favorable or not depending on the scene. The major disadvantages are the same as for black and white, namely a tendency toward excessive contrast and a potential for loss near the tails of the distribution. The Gaussian stretch generally produces a more equal treatment of each section of the histogram, with consequently less saturation. The mean of each component image should be assigned to mid-gray or higher. Figure 12 presents chromaticity diagrams before and after stretching of a color composite Landsat scene. A uniform distribution stretch was used.

Occasionally, more exotic techniques are useful; sinusoidal stretching (Ref. 16) is an example. This technique applies an intensity transformation defined by a multicycle sine function (so that for increasing input values, output brightnesses rise to maximum, fall off, rise again, and so forth). For color compositing, usually a different number of cycles is used for stretching each color component. This method can be useful whenever contouring (resembling that produced by left bit shifting for black and white images) is desired, or the user is interested in emphasizing subtle differences. Major drawbacks are that large scale feature boundaries tend to be lost in the maze of color, and the same color in two areas of the output image can result from different combinations of the input components, so that analogies cannot be drawn easily.

In some cases, a more controlled form of zonal stretching produces good results. For certain Landsat scenes of Iceland, Abrams (Ref. 17) has observed that the histogram of each spectral band is trimodal, thus providing effective separation between water, open ground, and areas of snow and ice cover. By carefully spreading each of these histogram zones to encompass the full brightness range in the output for each of the three color components, good color differentiation with color consistency is obtained within each zone (Ref. 13).

Color Filtering

Some filtering processes, such as MTF restoration filtering, can be conducted on the individual component images of a color composite with good results. In other cases, however, color filtering presents problems. Large scale brightness variations which mask local detail can be a problem in color pictures just as with black and white. The standard monochromatic solution, i.e., low frequency notch filtering, in particular subtractive box filtering, cannot be applied separately to the three components without encountering two major problems: loss of color information and a new artifact, color ringing.

The reason for loss of color is easy to visualize in the case of box filtering. After box filtering, the average brightness of any region of the image whose size is on the order of the filter size will be the midgray DC bias value. Thus any region of color in a composite of filtered images will be restricted in scale to less than the filter size. For smaller filter sizes, the resulting image can appear gray overall, with small local flecks of color. Many of these small color zones are due to the second problem, color ringing. This

artifact occurs when, due to a color boundary, there is a strong brightness transition in one or two but not all of the components. After filtering, ringing becomes evident in these components, but not in those without the transition. The appearance of the artifact is a zone of heightened saturation bordered by an induced area of the complementary color.

Both of these problems can be solved by observing that only the intensity constituent of the image requires filtering; it would be best to preserve color as much as possible. This objective can be attained (Ref. 18) by transforming the color image to either the chromaticity coordinates or the hue, saturation and intensity coordinates described earlier. Color information is isolated from intensity information, which is then filtered separately. A reverse transformation is then performed before display.

The principal advantage of the H,S,I system is that the two color components are important perceptually. Thus operations can be performed on the hue and saturation images during the same step, with reasonably predictable results. The chief advantage of the chromaticity approach is that it is computationally faster, since the actual transformation can in fact be avoided. Letting the subscript D represent a processed component ready for display and the subscript F represent a filtered component, then from Eq. (4):

$$B_D = bI_F = \left(\frac{B}{I}\right)I_F = B\left(\frac{I_F}{I}\right)$$

$$G_D = G\left(\frac{I_F}{I}\right) \qquad\qquad (5)$$

$$R_D = R\left(\frac{I_F}{I}\right)$$

(where image operations are performed pixel by pixel). Thus each of the original tristimulus components need only be multiplied by the coefficient image (I_F/I). In practice at IPL, where there exists a picture ratioing program capable of automatic scaling, the operations are usually performed in the order

$$I_R = \frac{I}{I_F}$$

$$B_D = \frac{B}{I_R}$$

$$G_D = \frac{G}{I_R} \qquad\qquad (6)$$

$$R_D = \frac{R}{I_R}$$

Another advantage of the chromaticity approach is that it can conveniently be extended to an arbitrary number of multispectral components, according to

$$M_D^{(1)} = M^{(1)}\left(\frac{I_F}{I}\right)$$
$$\vdots$$

$$M_I^{(n)} = M_I^{(n)}\left(\frac{I_F}{I}\right) \qquad (7)$$

$$I = \frac{1}{n}\sum_{i=1}^{n} M^{(i)}$$

After filtering, the individual components can then be displayed in any desired combination.

CLASSIFICATION AND MULTISPECTRAL TRANSFORMATIONS

The preceding sections have all dealt with processing techniques designed to solve certain classes of image analysis and display problems. The most important criterion in evaluating any image processing effort, however, is its success in better defining that information in the scene which is most important and meaningful to the user. Thus, considerable study has been devoted at JPL to determining which methods are most effective in extracting and displaying information of geologic significance from multispectral images, particularly with regard to lithologic mapping.

The techniques of mathematical pattern recognition naturally come to mind as a potential approach to categorizing surface materials based upon multispectral observations. Actual classification is one possible application of pattern recognition methodology. This can be implemented either by searching for naturally occurring groupings (i.e., cluster analysis) or by differentiating among different units based upon statistical properties observed within a known subset of the data (i.e., training). An alternative approach is to use pattern recognition or feature selection techniques to determine which subset of the multispectral data (including combinations of the original spectral bands) is most useful in isolating the groups of interest. These most significant component images can then be color composited for analysis.

Classification

The objective of automated classification as applied to multispectral images is to group observed materials into meaningful categories based on similarities in their spectral response characteristics. The n spectral bands comprising the data can be thought of as defining an n-dimensional cartesian space. Each element in the scene then corresponds to the point in the n-space whose coordinates are given by the brightness values at that pixel in each of the component spectral images. Thus the Landsat MSS, which acquires four spectral images, defines a four-dimensional pattern space. Classification is accomplished by partitioning this n-space into several disjoint regions, one for each class. Unknown points are classified by determining which region they occupy.

One approach to classification is to use the unknown data itself (actually usually a subset of the data) to define the naturally occurring groupings within the data. This approach is termed unsupervised classification, since predefined differences in certain specified characteristics do not govern the grouping. A variety of tech-

niques are available (Ref. 19) to find the naturally occurring clusters of data points in n-dimensional pattern space.

A different approach, called supervised classification, consists of using a set of "prototype" points, whose correct classification is known, to define decision boundaries for classifying other unknown points. A variety of supervised techniques are available (Ref. 20). One technique used at JPL is the Bayesian maximum likelihood algorithm, which classifies an unknown point $\underline{x}$ in that class S_K such that

$$h_K P(\underline{x}/S_K)\, P(S_K) \qquad (8)$$

is maximized, where $P(S_K)$ is a known a priori probability for the occurrence of class S_K, and h_K is a weighting factor allowing correct classification in certain classes to be emphasized relative to others. The probability $P(\underline{x}/S_K)$ that the unknown point is in class S_K is usually assigned a Gaussian distribution

$$P(x/S_K) = \frac{1}{(2\pi)^{\frac{n}{2}}|\phi_K|^{\frac{1}{2}}} \cdot$$

$$\exp\left[-\frac{1}{2}(\underline{x}-\underline{\mu}_K)^t \phi_K^{-1}(\underline{x}-\underline{\mu}_K)\right] \qquad (9)$$

where the class means $\underline{\mu}_K$ and covariance matrices ϕ_K are determined by training using the known prototype points. In a variation of this approach, Addington (Ref. 21) has devised a hybrid classifier which combines a parallelepiped partitioning algorithm with the Bayesian decision rule. The parallelepiped algorithm is applied first, and its results actually used to eliminate unlikely alternatives. Many applications of the Bayes algorithm are thus eliminated, with an appreciable savings in computer time.

These classification techniques have been applied to a test site in Northern Arizona, using Landsat multispectral images, to assess their value in lithologic mapping (Ref. 19). Results were unsatisfactory: only about 50 percent of the study area was mapped correctly. The primary reasons for this failure are the gradational and non-homogeneous nature of geologic units, and the presence of vegetation. Another important reason is that, for the geologic units involved, there simply is not much variation in that part of their spectral response characteristics detected by the Landsat bands.

Feature Selection

A somewhat different approach to information extraction is to use the methods of pattern recognition and feature selection to determine which spectral bands, or which functions of spectral bands provide the best separation of classes. These bands or functions can then be enhanced and color composited for display. The trained eye of the geologist can then still be used to take advantage of spatial correlation and contextual relationships.

One useful set of procedures is described by Siegal (Refs. 19 and 22). One of these techniques

employs a multiple group discriminant analysis procedure to determine, in a stepwise manner, which dimensions in pattern space are the most effective in promoting separability. Each added dimension is evaluated in terms of the ratio of between group and within group dispersion (variances) in the resulting subspace with the objective of maximizing this ratio. This technique can be applied, for example, to indicate which three of the six Landsat ratio images (4/5, 4/6, 4/7, 5/6, 5/7, 6/7) would be best for discriminating between certain classes of material. The method has also been applied to field acquired spectral reflectance data in the wavelength region 0.45 to 2.4 microns to determine which spectral bands within this region would be most effective in geologic mapping (Ref. 23); and in fact, bands substantially superior to the Landsat bands have been found. Another of the procedures first normalizes each dimension in pattern space to equalize observed dispersion, then performs a canonical transformation (principal component transformation) to find which linear combinations of the original coordinates produce the greatest separation of the groups. The first three principal components can then be displayed as a color composite.

Principal component analysis has also been applied to images without using training areas. In this case, the entire image is treated as belonging to a single class and a covariance matrix $C = [c_{ij}]$ describing the second order multivariate statistics is computed according to

$$c_{ij} = \frac{1}{N} \sum_{k=1}^{N} (x_i^{(k)} - \mu_i)(x_j^{(k)} - \mu_j) \qquad (10)$$

where $x_i^{(k)}$ is the kth pixel of the ith component image, and μ_i is the mean of the ith component. Then the equation

$$KC = \underline{\lambda}K \qquad (11)$$

is solved to obtain the eigenvectors k_ℓ and eigenvalues λ_ℓ of C (where ranges over the n spectral components). The resulting eigenvector matrix K is then used to transform pattern space; this transformation is also known as a Karhunen-Loeve transformation (Ref. 24). Effectively the processes KCK^T diagonalizes the covariance matrix, thus removing the correlation which existed among the original spectral bands. Intuitively, it seems reasonable to expect (Ref. 25) that being uncorrelated, the first three principal components (those with the largest and hence the greatest variance) should contain more information than any three of the n original components. Since the process is not supervised by use of known prototypes from different classes, there is no guarantee that better separation between materials of interest will be present in a principal component color display. In one interesting variation, Fontanel, et al. (Ref. 26) have performed a classification, then a principal component analysis applied separately (in a zonal fashion) to that spatial region defined by each class, to achieve an enhancement of information within that class.

An additional application of principal component analysis is in evaluation of signal-to-noise ratios. Since the Karhunen-Loeve transformation is unitary, total variance is preserved (Ref. 27). Visual analysis suggests that the last principal component (component 4 in the case of Landsat) is often almost entirely noise. Thus the variance of one of the original components can be divided by the eigenvalue of the last principal component to give a measure of the signal-to-noise ratio in the original band. The square root of such a value was quoted earlier in this paper.

ACKNOWLEDGEMENTS

REFERENCES

1. Goetz, A.F.H., "Use of Landsat Imagery for Geological Analysis," in these Proceedings.

2. Kahle, A. B., "Thermal Inertia Mapping Using Remote Sensing," in these Proceedings.

3. Gillespie, A. R., and Kahle, A. B., "The Construction and Interpretation of a Digital Thermal Inertia Image," submitted for publication.

4. Goetz, A.F.H., Billingsley, F. C., Gillespie, A. R., Abrams, M. J., Squires, R. L., Shoemaker, E. M., Lucchitta, I., and Elston, D. P., "Application of ERTS Images and Image Processing to Regional Geologic Problems and Geologic Mapping in Northern Arizona," Jet Propulsion Laboratory Technical Report 32-1597, 1975.

5. Bernstein, R., "Digital Image Processing of Earth Observation Sensor Data," IBM Journal of Research and Development, Vol. 20, No. 1, pp. 40-57, January 1976.

6. Bernstein, R., "Scene Correction (Precision Processing) of ERTS Sensor Data Using Digital Image Processing Techniques," Proceedings of the Third Earth Resources Technology Satellite-1 Symposium, Vol. 1, Section B, NASA SP-351, pp. 1909-1927, December 10-14, 1973.

7. Earth Resources Technology Satellite Data Users Handbook, Goddard Space Flight Center, Greenbelt, MD, 1972.

8. Elliott, D. A., "Digital Cartographic Projection," in these Proceedings.

9. Bracewell, R., The Fourier Transform and Its Applications, McGraw-Hill, 1965.

Figure 1a. Landsat scene of Goldfield, Nevada, with six line striping.

Figure 1b. Striping has been reduced by forcing sensor distributions to match.
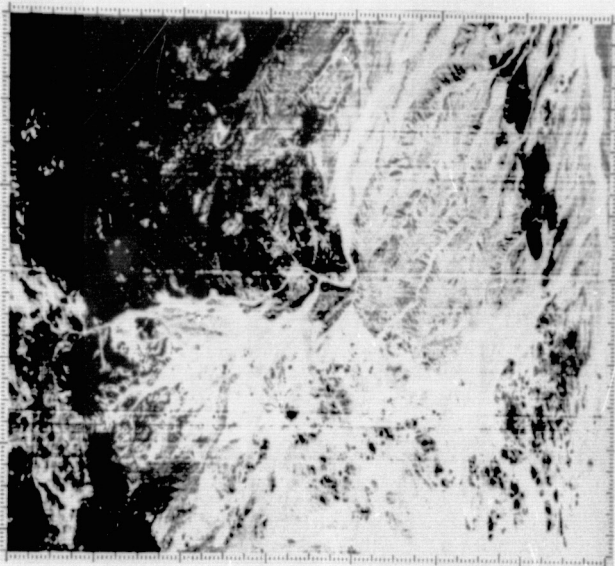
Figure 2a. An aircraft scanner image with stripe
noise.
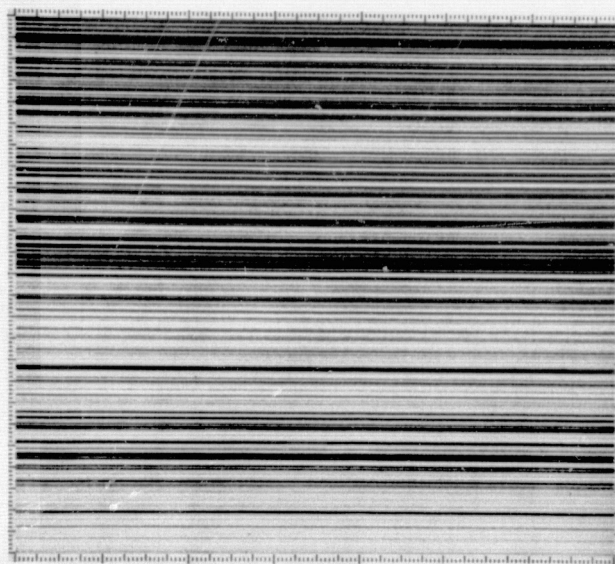
Figure 2b. An additive approximation to the stripe
noise can be obtained by vertical high
pass filtering of line averages. The
resulting noise component shown here has
been contrast enhanced for display.
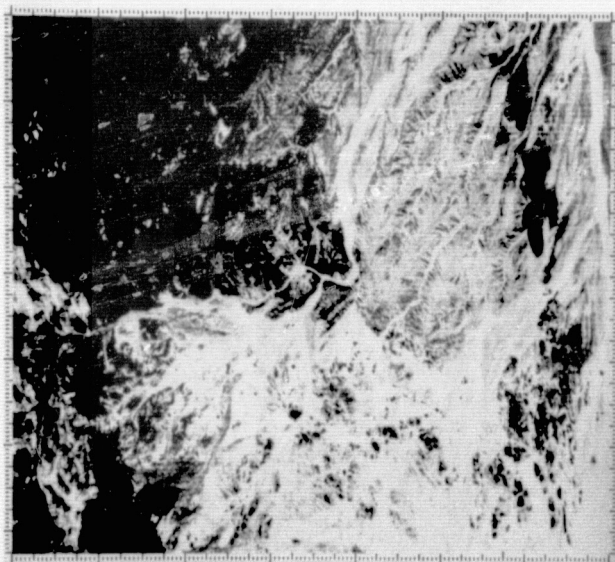
Figure 2c. The scanner image with noise subtracted.
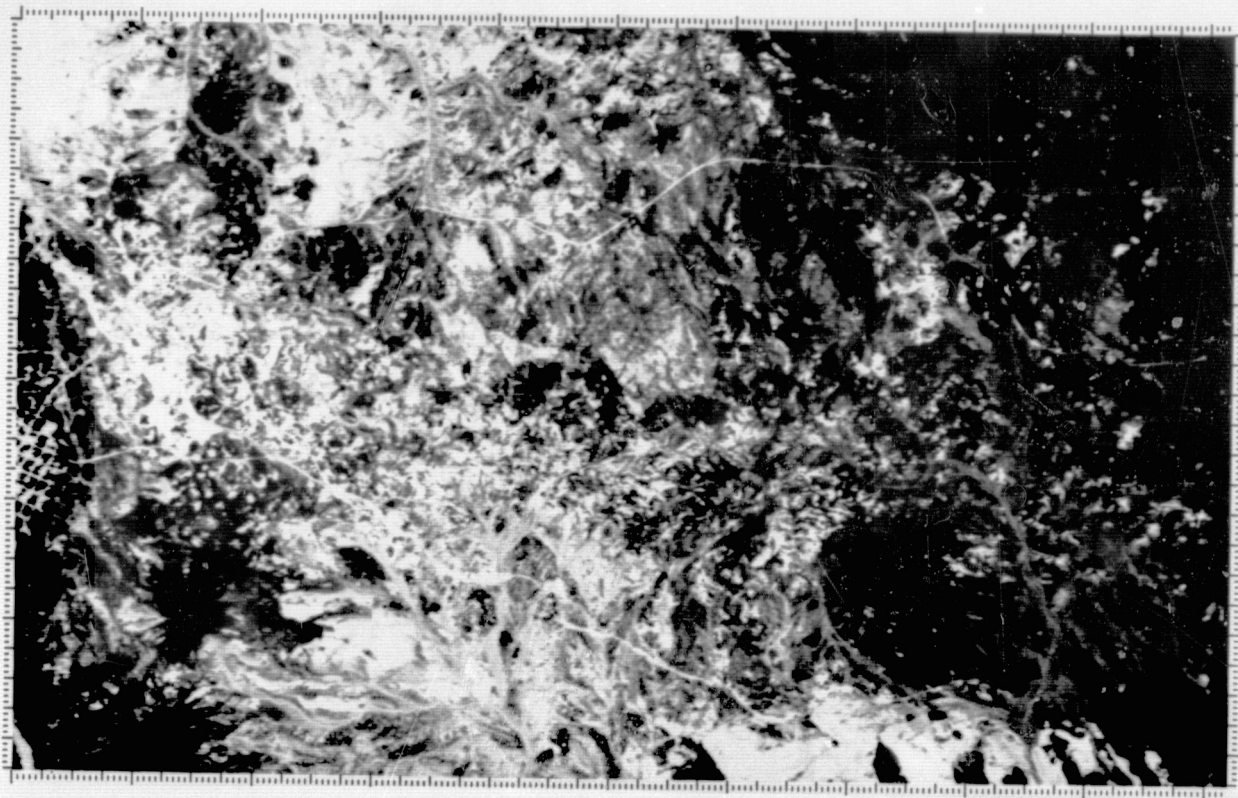
Figure 3a. An aircraft scanner image of Goldfield, Nevada.
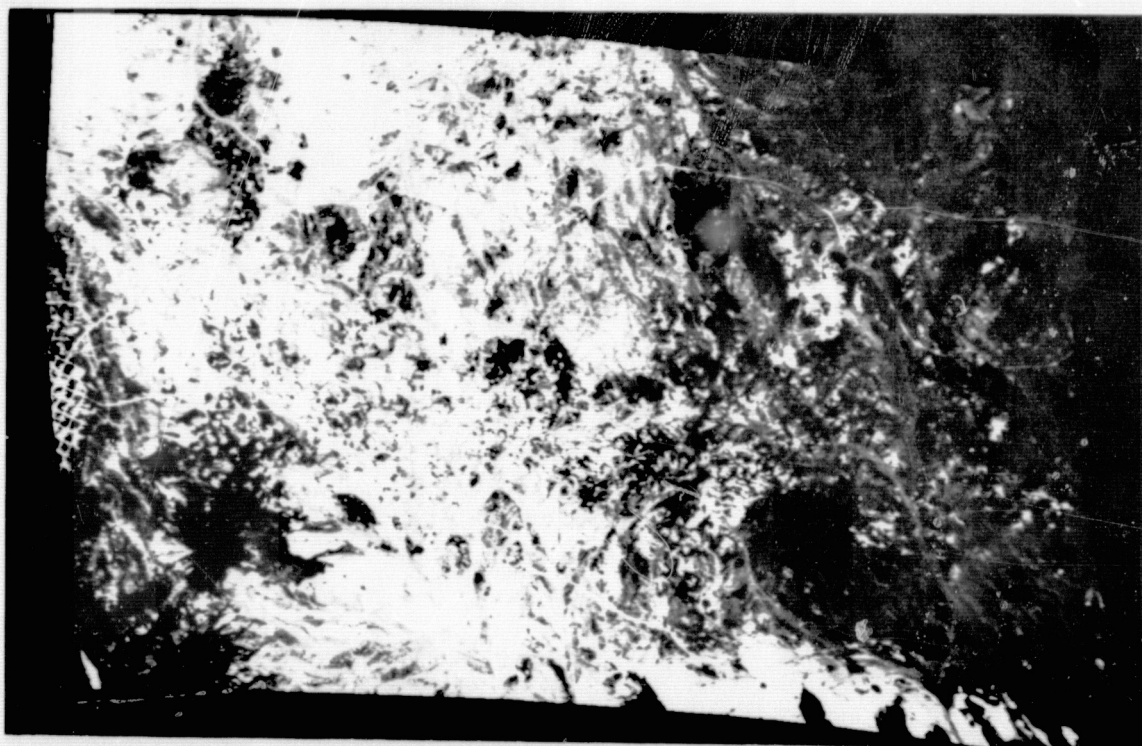
Figure 3b. The Goldfield image after geometric transformation to a Universal Transverse Mercator projection.
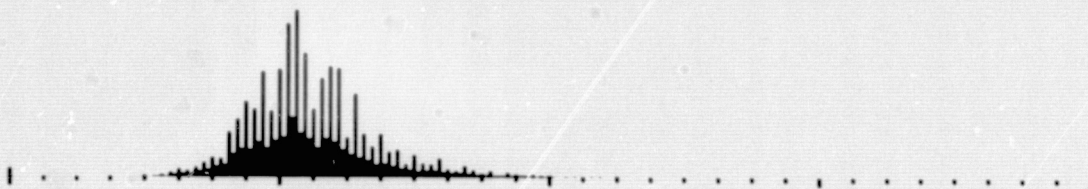
Figure 4a. Histogram of the original Landsat scene of Figure 5a.
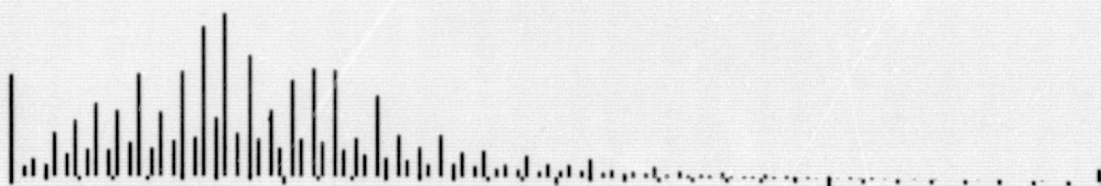
Figure 4b. Histogram of the same scene after an automatic percentage saturation contrast stretch (image shown in Figure 5b).
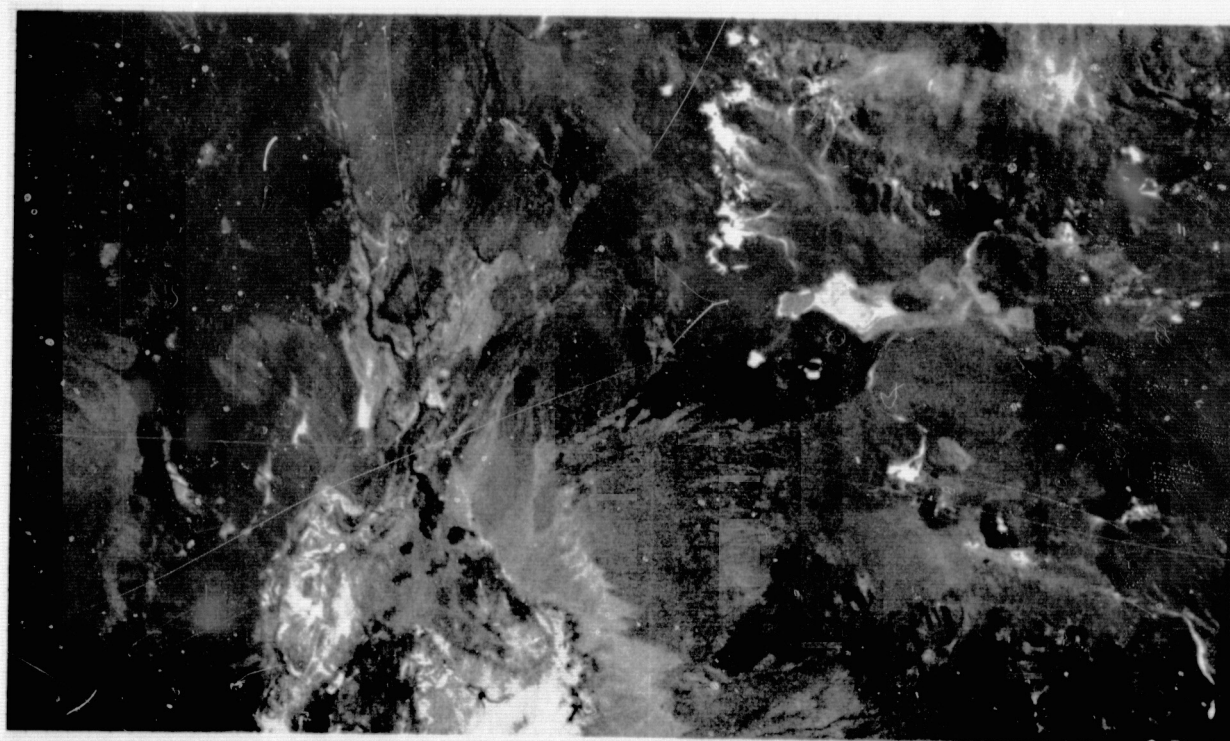
Figure 5a. A Landsat scene of the Andes at the border of Chile and Bolivia, shown with no contrast enhancement. Contrast is insufficient.
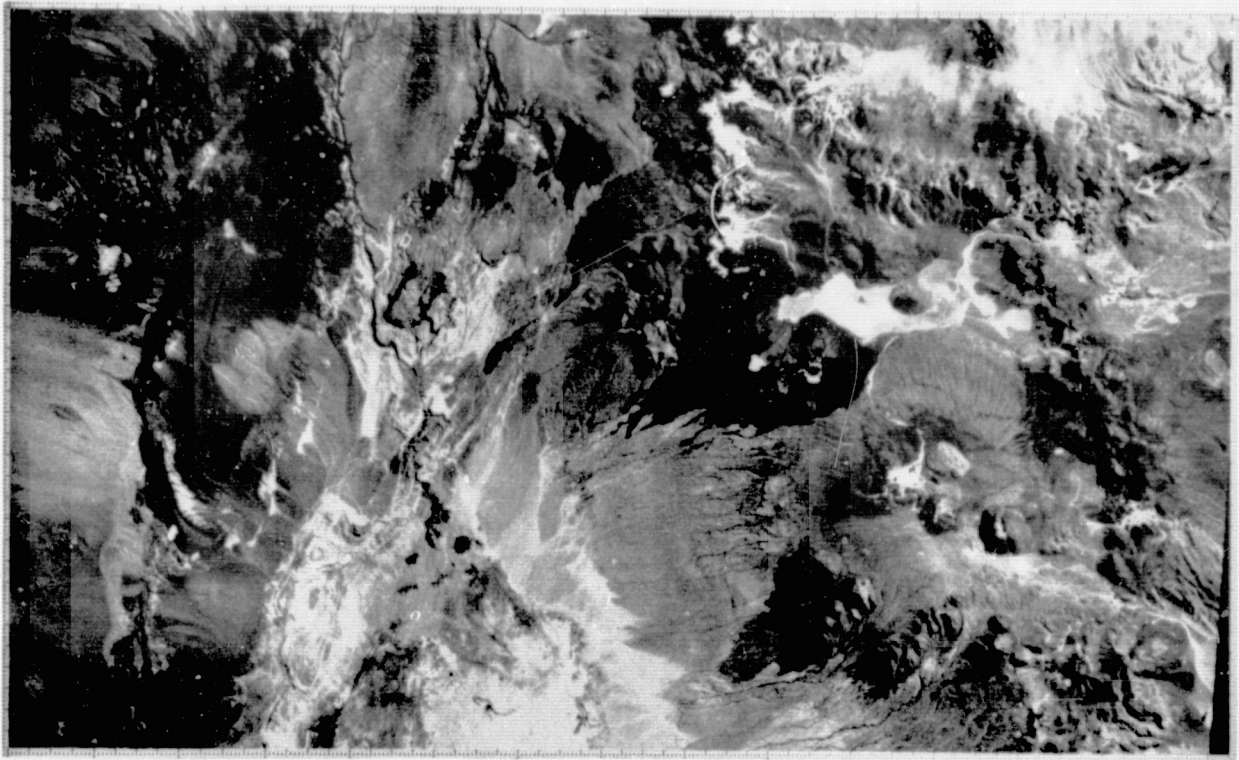
Figure 5b. The same scene after an automatic linear contrast enhancement in which 4 percent of the histogram is saturated.
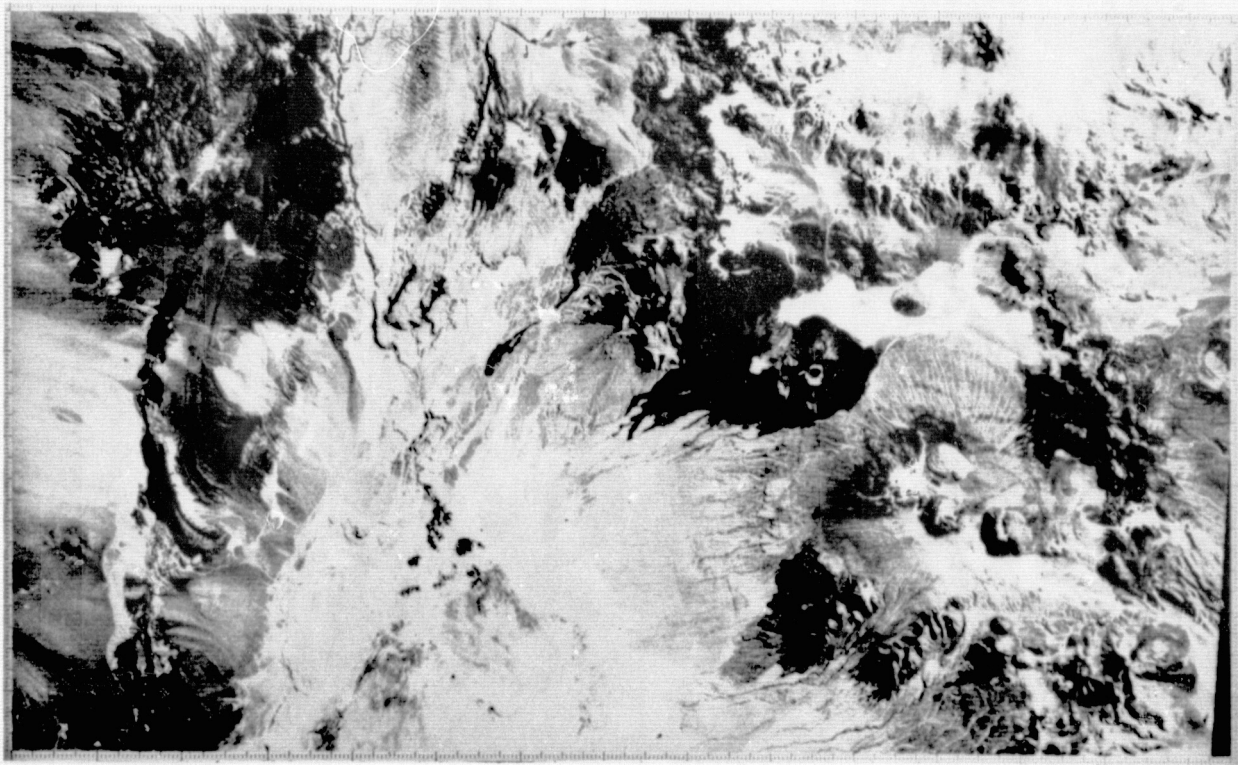
Figure 5c. The Chile-Bolivia scene after a uniform distribution automatic stretch. In this scene, detail at the extremes of the histogram is lost, such as around the volcano to the right of center.
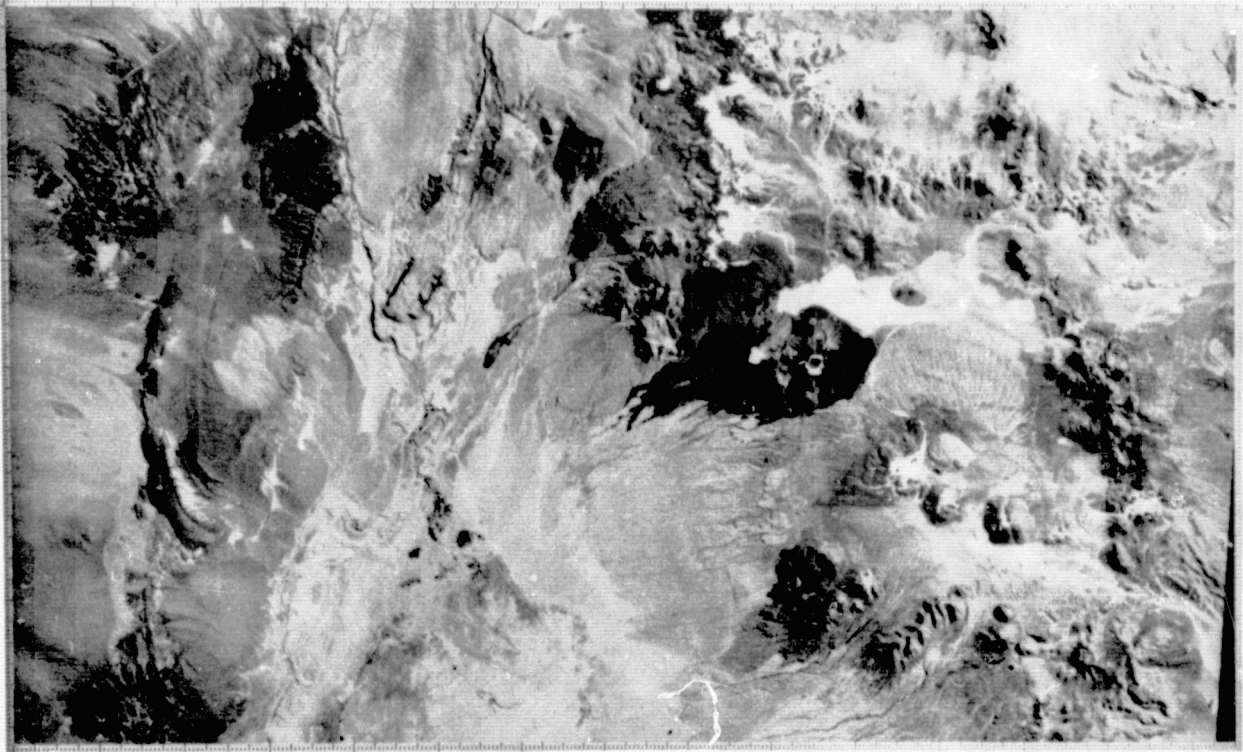
Figure 5d. A Gaussian stretch produces a better display near the tails of the histogram, but is weaker for midgray values.

Figure 6a. An aircraft scanner image of Patagonia, Arizona, with coherent sensor noise.

Figure 6b. A vertical power spectrum of the Patagonia frame.

Figure 6c. The scanner image after vertical Fourier domain notch filtering to remove noise spikes.
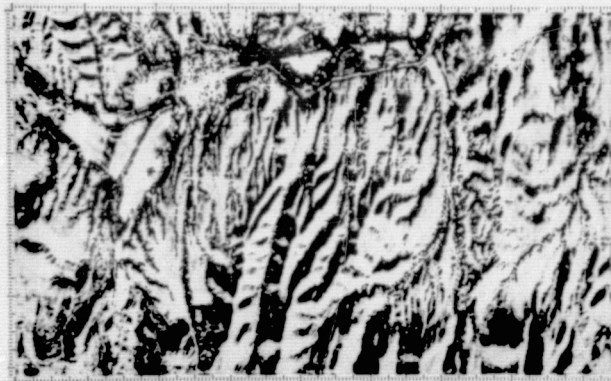
Figure 7. The Patagonia image of Figure 6 after high pass filtering to remove shading.
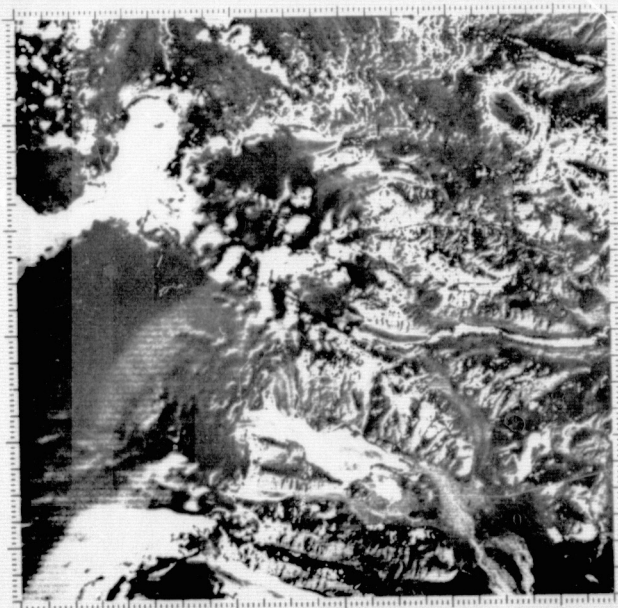
Figure 8a. A portion of a Landsat scene of Iceland. Apparent wraparound is due to zonal stretch.
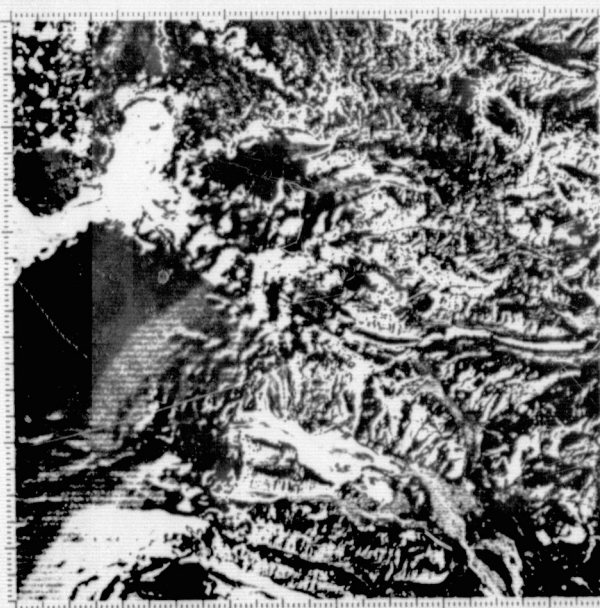
Figure 8b. The same image after "pseudo-MTF" filtering.

Figure 9a. Artifacts can be a problem with box filtering. Here a one line, horizontal filter preferentially enhances linear feature at a small angle to the horizontal.
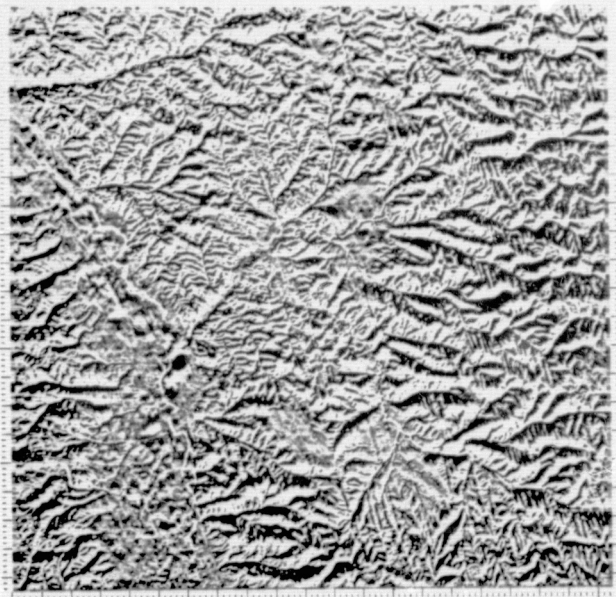
Figure 9b. In a one sample, vertical filter, linear features oriented at the same angle to the vertical are emphasized.

Figure 9c. Preferential directional enhancement is minimized with a square filter.
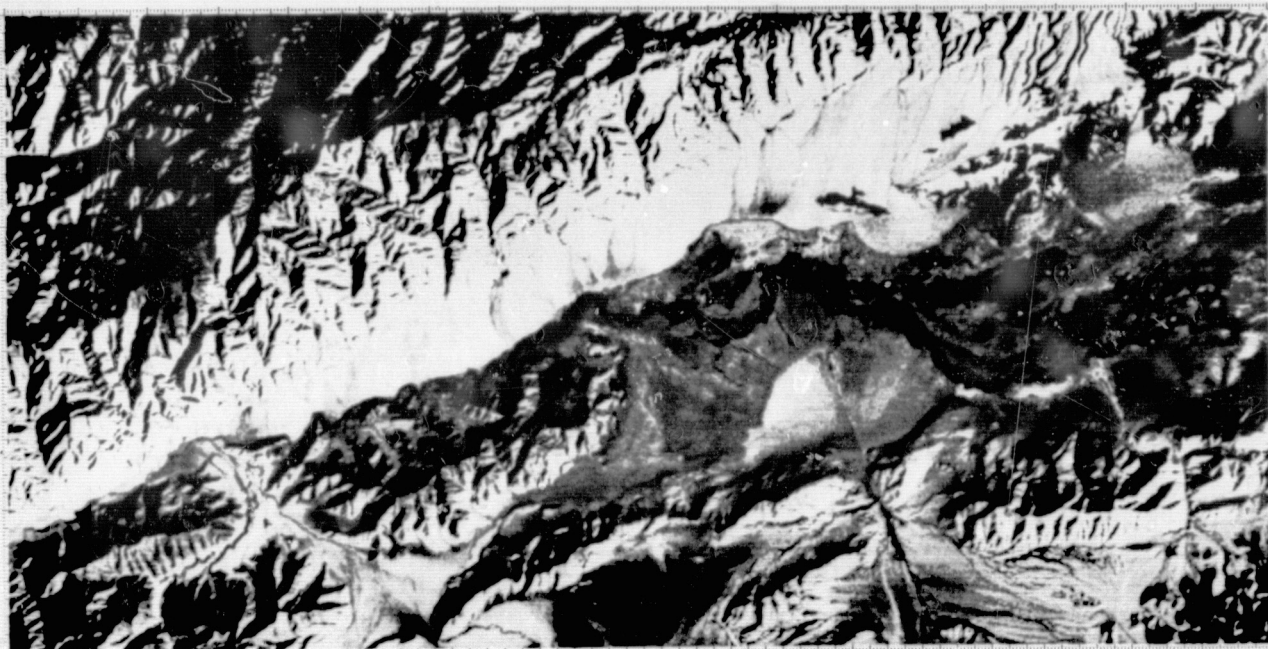
Figure 10a. A Landsat image of the Altyn-Tagh fault zone in China.
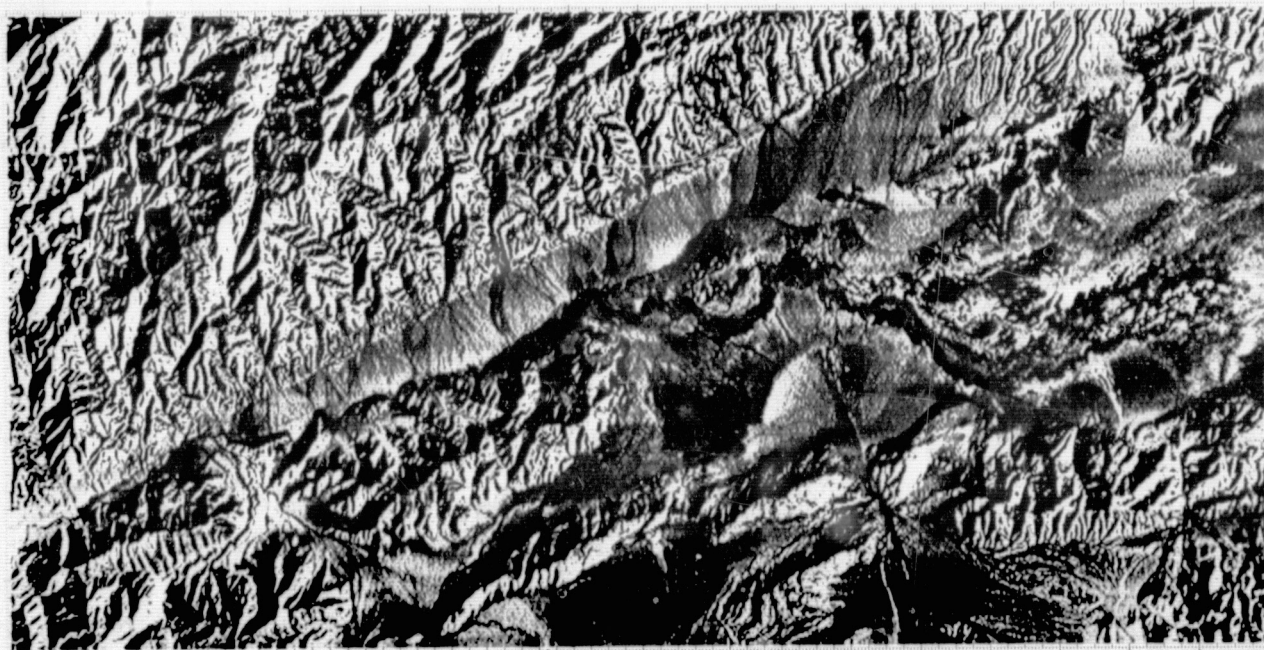
Figure 10b. The Altyn-Tagh frame after low frequency notch filtering and high frequency enhancement to emphasize local detail.
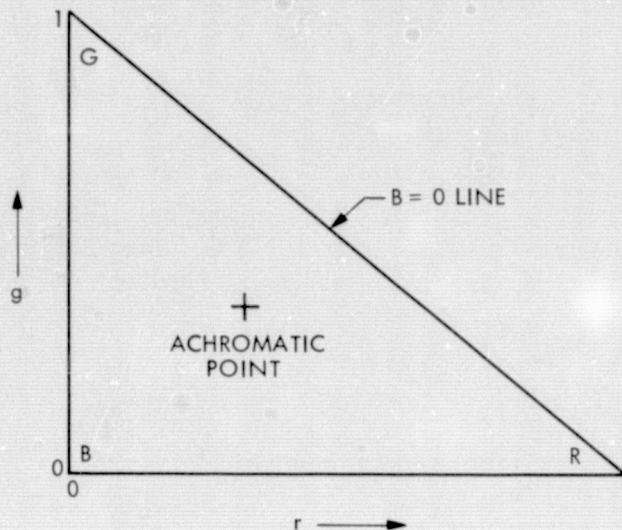
Figure 11. The format of a chromaticity diagram. Saturated blue, green and red correspond to the regions marked B, G and R, respectively. The region around the achromatic point represents less saturated colors.
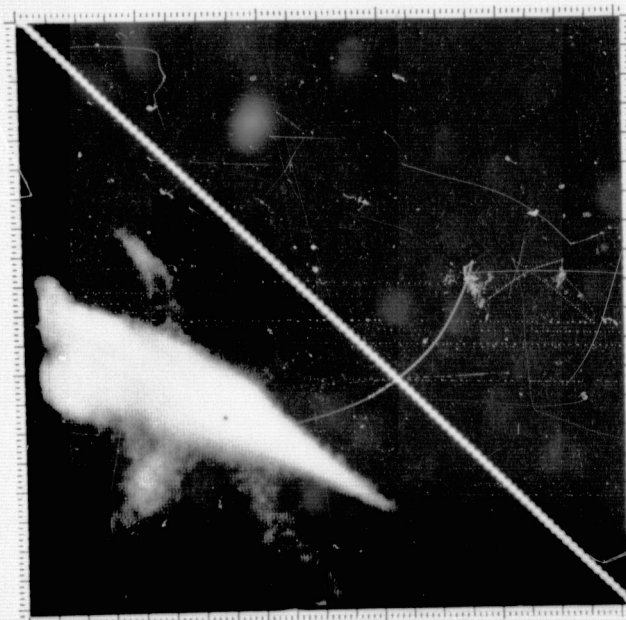
Figure 12a. Original chromaticity distribution of a Landsat scene.
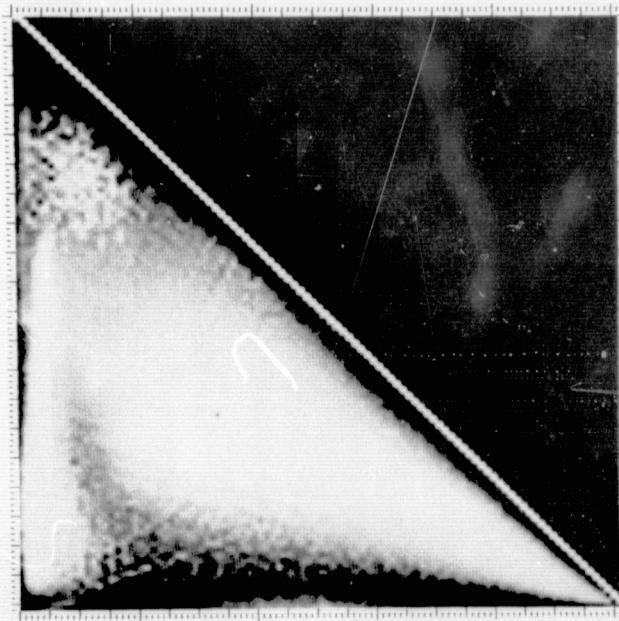
Figure 12b. Chromaticity distribution of the same scene after contrast enhancement.
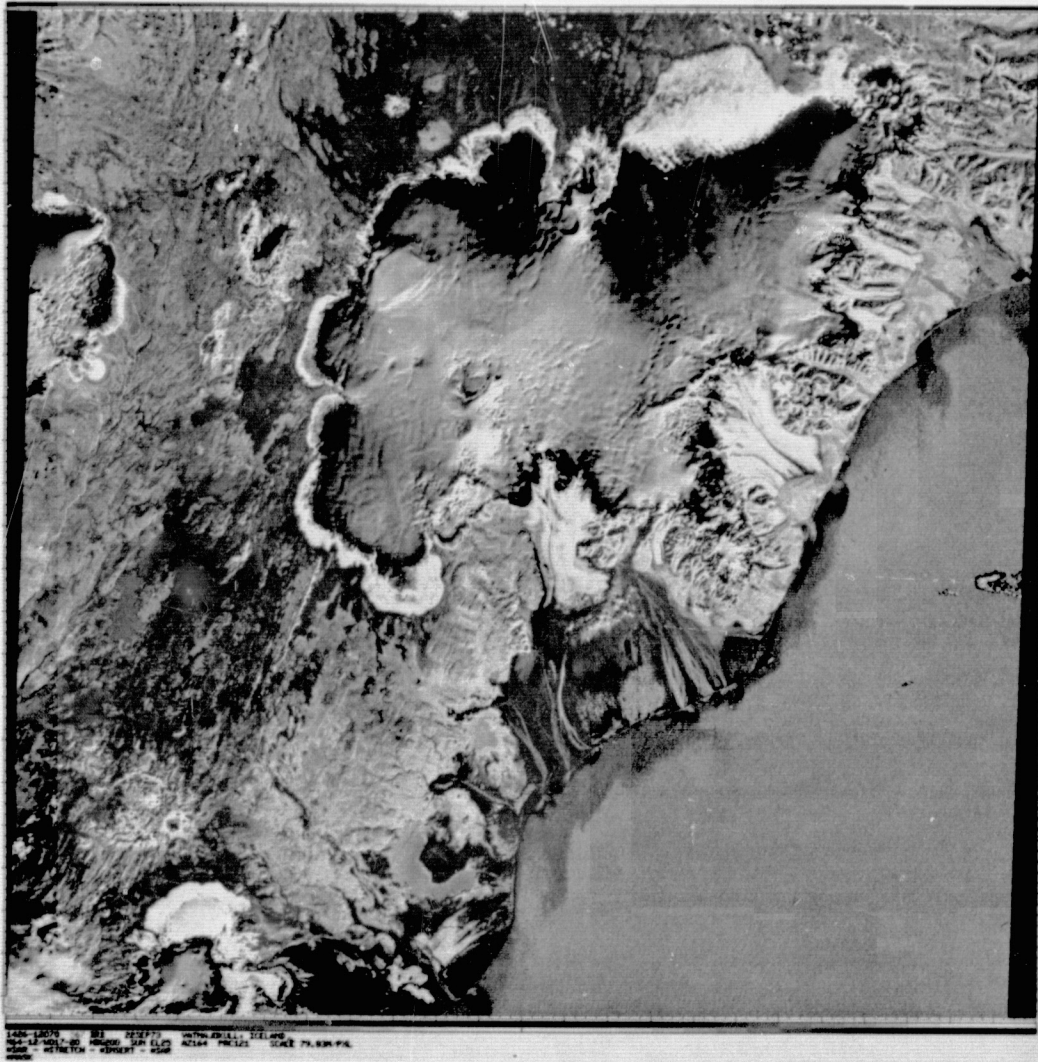
Figure 13. Zonal stretching applied to a Landsat scene of Iceland. The original histograms are trimodal, allowing areas of snow, open ground, and water to be treated separately. Each zone was stretched in each of the three component images to encompass the full brightness range. Colors are consistent throughout the image within each of the three zones.

DIGITAL CARTOGRAPHIC PROJECTION

Denis A. Elliott

Jet Propulsion Laboratory
Pasadena, California

ABSTRACT

A general technique for digital cartographic projection of planetary images taken from spacecraft is described. Several applications of map projections are illustrated. Also, some specific problems which arise at one phase of projection are mentioned, and their solutions as implemented at the Image Processing Laboratory of the Jet Propulsion Laboratory are summarized.

I. INTRODUCTION AND ACKNOWLEDGEMENTS

This paper is intended to familiarize the reader with a technique for cartographically projecting digital images of planets or satellites which has been developed over a period of several years at the Image Processing Laboratory (IPL) of the Jet Propulsion Laboratory, Pasadena, CA. It assumes that the reader has some previous experience in the field of digital image processing.

Although this author has contributed to the effort to implement projection capability at IPL, others have made more fundamental and extensive contributions. James Soha and Alan Gillespie developed many of the algorithms now in use. C. A. Rofer and A. A. Schwartz extended their work considerably. H. J. Frieden and J. B. Seidman developed the techniques used to actually geometrically transform the input picture into the output picture once the proper transformation has been specified. John Kreznar developed the method used to remove geometric distortions prior to projection, which is discussed in Section VI.

A program has been written at IPL to perform cartographic projections. This program has the name 'MAP2'. It runs under control of the VICAR system developed at IPL. MAP2 will frequently be referred to throughout this paper. In fact, this paper is no more than a brief description of the program as it currently operates.

II. MAP PROJECTIONS AND THE REASONS FOR USING THEM

A map projection is an attempt to represent the curved surface of a sphere (or, more generally, an oblate spheroid) on a flat plane. Any such projection involves distortion of distances between points and of angles between lines joining widely separated points. Some projections preserve shapes of small areas, others preserve areas of small zones, some preserve neither. It is impossible to preserve both area and shape, even for small zones. There are several reasons for taking an input picture and distorting it so that the features in the output picture fall where they would if the picture were actually a map of the area, made using some standard map projection such as a Mercator or Polar Stereographic Projection. But all reasons are related to one fact of overriding importance: the input picture is itself a representation of a sphere on a plane and therefore has its own particular distortions built into it. It is frequently advantageous to trade one set of distortions for another.

The most common reason for map projecting a picture is to facilitate mosaicing of several pictures of adjacent areas. Since these frames are taken from a spacecraft in orbit around or flying by a planet, each one is taken from a somewhat different angle and distance. Thus the distortions of shape and scale in adjacent pictures may be so different that the set of pictures cannot be mosaiced without severe misalignments at picture boundaries. Map projections are continuous representations of the surface, so a set of frames projected using the same parameters will mosaic perfectly. If the photometric accuracy and the geometric accuracy are both high then, in a computer-made mosaic, picture boundaries are difficult or impossible to detect. Figure 1 is a mosaic made at IPL by Joel Mosher from Mariner 10 Mercury pictures. It consists of 76 separate frames. Careful examination of the negative or good prints reveals a few boundaries which became apparent because of small projection errors caused by uncertainties in the camera pointing (see section VI for a discussion of the importance of accurate knowledge of camera pointing). But for the most part the final product can only be described as magnificent. The mosaic covers the area from 0° to 90°W longitude, -20° to -70° latitude. Mosaics such as the one in Figure 1 can be used as the basis for mapmaking, and in fact similar techniques to those discussed in this paper have been used by various mapmakers since the advent of airplanes.

Even when mosaicing is not envisioned, map projection can be a useful image enhancement tool. Pictures taken along a line of sight not close to the normal to the surface suffer from severe foreshortening of features. Such shape distortion can cause severe problems if the morphology of craters is being studied, to give one example. An orthographic projection centered on a point within the field of view will eliminate foreshortening and still preserve much of the visual appearance of the scene.

For many purposes it is desirable to know the latitude and longitude of points in the picture. In an unprojected frame the latitude and longitude are complicated functions of many variables, and the values of many of the variables will be unknown to most users. On the other hand, in a map

---

projection the coordinates are relatively simple functions of a few variables whose values can easily be displayed in the picture's label.

Another interesting use of map projection is to display global features in such a way that their true nature becomes apparent. For example, Mariner 10 pictures of Venus' atmosphere show a series of prominent bands approximately parallel to the equator (Figure 2). As a first step in understanding these bands it is important to know if they are truly parallel to latitude circles or if they have a meridianal component of motion. In a polar stereographic or orthographic projection, latitude circles project as circles. Figure 3 is such a projection centered at the south pole of Figure 2. It shows that, at least near the pole, the bands are not circular.

III. OVERVIEW OF THE PROJECTION TECHNIQUE USED BY MAP2

The entire process of cartographic projection can be summed up very briefly as: "Take each point in the input picture, determine its latitude and longitude, then determine its corresponding position in the output picture using the equations which describe the projection being used. Finally, actually put the output picture together from the input." The preceding description is valid in the sense that the net result of following it will produce (in the ideal case) properly projected pictures. However, the specified sequence of events turns out to be impractical for two reasons. One reason is related to problems in the actual construction of the output picture. The other is related to an instability in the computation of latitude and longitude from position in the input picture.

The actual production of the output picture turns out to be by far the most difficult and time-consuming part of the projection process. Depending on details of the camera viewing geometry and the desired projection, densely crowded points in the input picture may map to sparsely spaced output points, or vice versa. Integral picture elements (pixels) will in general map to non-integral coordinates in the output picture. The specification of the exact transformation of each input pixel would require prohibitive amounts of computer time. For example, Viking Orbiter frames contain over 1 1/4 million pixels. H. Frieden and J. Seidman at IPL have worked extensively on the problem of "rubber-sheet stretching" pictures and have developed several programs for performing such transformations. The basic procedure is to define a set of rectangles or general quadrilaterals in the output picture. The exact transformation from input to output is specified only at the vertices. Within each quadrilateral the mapping is interpolated from its values at the vertices. The problem of an output pixel corresponding to a fractional input pixel (non-integral line and/or sample) is handled by bi-linear interpolation from the four nearest neighbors in the input picture. See Reference 1 for more details. The MAP2 program fetches one of the rubber-sheet stretch programs after computing the mapping transformation at a set of "tiepoints" in the output picture. The particular program fetched requires that the output tiepoints form a rectangular array. This forces us to work backwards in a

sense. A set of arbitrarily selected input points will not map to a rectangular array in the output frame, so we begin by selecting a rectangular array in the output picture, using the projection equations to determine the latitude and longitude of the points, and then computing their line and sample coordinates in the input frame. The specification of line and sample coordinates of the same points in both pictures allows the geometric stretch program to actually produce the final product.

The summary of digital cartographic projection which began this section can now be restated accurately as: "Specify a rectangular grid in the as yet not produced output frame. Use the equations which describe the desired projection to determine latitudes and longitudes of the selected points. Use a knowledge of the camera's viewing geometry to determine the line and sample coordinates of these same points in the input picture. Fetch the geometric transformation program to produce the output frame."

The above order of computation has another important advantage over the "natural" order originally given. The computation of line and sample coordinates from latitude and longitude is much less sensitive to small errors in input data than the inverse computation of latitude and longitude from line and sample in the input picture. But the two computations in the output picture do not differ noticeably in sensitivity, at least for the projections presently implemented at IPL. These are Mercator, Two-Standard Lambert Conformal Conic, and oblique and polar stereographic and orthographic. Thus the actually used technique is not seriously hampered by floating point precision problems in a computer. The other, seemingly more natural, approach involves at one point the subtraction of two nearly equal numbers with a typical loss of four significant figures.

Map projection thus divides neatly into three phases: determination of the latitude and longitude of a set of regularly spaced points in the output picture; determination of the line and sample coordinates of the corresponding points in the input picture; the actual transformation of input into output (a "rubber sheet" stretch). The details of the third phase will not be presented here. They are covered in Reference 1.

IV. MAP2 OPERATION - PHASE 1

Let λ = E. longitude of a given surface point;

ϕ = latitude of the same point;

x = sample number of a given picture element, increasing to the right,

y = line number of a given picture element, increasing down.

The coordinates of a point on the surface of the planet in the output (projected) picture are given by

1) (x,y)out $= \vec{F}_m (\lambda, \phi)$,

where $\vec{F}_m$ is the vector-valued function, describing the desired projection. For example, the equations for a Mercator projection of a sphere of radius R km. with north at the top are:

2)

$$(\vec{F}_m)x = \frac{\pi}{180} \frac{R}{S} (\lambda - \lambda_0) + X_0$$

$$(\vec{F}_m)y = \frac{\pi}{180} \frac{R}{S} Ln\left[tan \left(45 + \frac{\phi}{2}\right)\right] + Y_0$$

where angles are measured in degrees, S is the scale of the projection in km/pixel at the equator, and the point on the equator at longitude λ_0 lies at (x_0, y_0) in the image.

In the first phase MAP2 picks 400 equally spaced points across the output picture and computes

3) $$(\lambda, \phi) = \vec{F}_m^{-1} (x,y)$$

for each point. $\vec{F}_m^{-1}$ is built into MAP2 for the Mercator and Two-Standard Lambert Conformal Conic projections for an oblate spheroid. More projections could easily be added to the list.

Phase 1 is the simplest step in the projection process. The only complication is that $\vec{F}_m$ inverse may not exist for some points in the output picture, since not all projections fill the entire plane. MAP2 handles this problem by specifying to the rubber sheet stretch program that the output point corresponds to a point far off the input image. In such cases the transformation program in phase 3 puts zeroes (black) in the output picture.

V. MAP2 OPERATION - PHASE 2

The input (unprojected) picture can be thought of as a set of points on the planet's surface whose coordinates satisfy the equation

4) $$(x,y)_{in} = \vec{F}_c (\lambda, \phi)$$

where $\vec{F}_c$ is a vector-valued function which describes the camera's viewing geometry. The exact nature of $\vec{F}_c$ depends on the optical design of the camera, the size and shape of the target planet, the position of the spacecraft, and the direction in which the camera is pointed. It should be clear by now that $\vec{F}_c$ must be known in order to project pictures. $\vec{F}_c$ is always defined for optical systems involving a lens or mirror and a true image plane. Of course, some points may map to coordinates on the image plane which are not in the field of view. Also two points of the planet will map to one point in the image. One will be on the front (visible) side and one on the back. Finally, there may be points in the input image which are not on the planet. This situation occurs when photographs of the planet's limb are taken, and for pictures taken from far away when the entire planet is in the field of view. Since we are computing $\vec{F}_c$ and not its inverse (as would have been the case were we using the "normal" order of computation mentioned in Section II) none of these complications poses any problem. Remember that if an output point corresponds to a point not in the camera field of view, the program fetched in phase 3 sets the brightness there to zero.

At IPL our past projection experience has involved pictures from Mariners 6, 7, 9, and 10 and the two Viking Orbiters. The optical design

of all the cameras on these missions was similar. A lens is used to form a true image which is electronically sampled. This means that the type of input pictures MAP2 has been designed to handle are true perspective projections, and the form of $\vec{F}_c$ is thus the same for all pictures. Map projections and mosaicing have many important applications for ERTS imagery (Earth Resources and Technology Satellite). But ERTS cameras are scanning cameras (each pixel obtained at a different time). Major changes to the structure of the MAP2 program may be required to handle the new form of $\vec{F}_c$ involved. Such changes are planned but implementation has not begun and this paper will confine itself to a discussion of $\vec{F}_c$ for Mariner and Viking cameras.

Figure 4, taken from an internal JPL report on the MAP2 program written by C. Rofer and A. Schwartz, illustrates the type of camera viewing geometry which concerns us. In the figure the point labeled "C" is actually the camera's lens. The part of the figure dealing with the image plane has been reflected about a line passing through C perpendicular to the camera line of sight. This has been done to clarify some of the geometric relationships between surface points and their corresponding image points. No mathematical relationships are disturbed by this reflection because of a fundamental theorem of optics which states that, as seen from the lens, the angular sizes of the image and the real object are identical. The image, of course, is actually to the right of C in the figure.

Three different Cartesian coordinate systems are involved in the computation of $(x,y)_{in}$ from (λ, ϕ). The first system will be called the planetary surface system. It rotates with the planet so that a point's coordinates in this system are fixed. Its origin is the center of the planet. The +z-axis points to the north pole; the +x-axis to the point $(0°, 0°)$; th +y-axis to the point $(90°, 0°)$. The planetary surface coordinates are computed from the longitude and geocentric latitude as follows:

$$x_p = R(\phi) \cos \lambda \cos \phi$$

5) $$y_p = R(\phi) \sin \lambda \cos \phi$$

$$z_p = R(\phi) \sin \phi$$

Equation 5 is valid on an oblate spheroid described by

6) $$\frac{x^2}{R_{eq}^2} + \frac{y^2}{R_{eq}^2} + \frac{z^2}{R_{pole}^2} = 1$$

On such a spheroid $R(\phi)$ is given by

7) $$R(\phi) = \frac{R_{eq} R_{pole}}{\sqrt{R_{pole}^2 \cos^2 + R_{eq}^2 \sin^2}}$$

The second important coordinate system is the camera coordinate system. Its origin is at the center of the camera lens, C, in Figure 4. Its +z-axis points along the line of sight towards the planet (into the picture, not out of it). Its

+x-axis is parallel to the increasing sample direction in the image plane, and its +y-axis is parallel to the increasing line direction. Note that both systems so far described are right-handed. If the vector $\overrightarrow{OP}$ from Figure 4 is known in the planet system and if the vector $\overrightarrow{OC}$ is also known in the planet system, then we have

8) $$\overrightarrow{CP}_p = \overrightarrow{OP}_p - \overrightarrow{OC}_p$$

The P-subscripts are there to indicate that the vectors are expressed in the planetary surface system. Define the OM matrix to be the rotation matrix which converts vectors from the planetary surface system to the camera system. Then

9) $$\overrightarrow{CP}_c = [OM]\ \overrightarrow{CP}_p$$

The third coordinate system is the image plane system which we have been implicitly using throughout. Its origin is the intersection of the camera line of sight and the image plane. Its +x-axis points towards increasing sample numbers (to the right). Its +y-axis points towards increasing line numbers (down). Note that this is a two-dimensional left-handed system. The coordinates of P^1, the image of P, in the image system are determined as follows. First, note that

10) $$\overrightarrow{CP}^1 = s\overrightarrow{CP}$$

where s is some scalar. It should be clear from Figure 4 that

11) $$s = \frac{(\overrightarrow{CP}^1)_z}{(\overrightarrow{CP})_z}$$

in the camera system. Therefore

12) $$s = \frac{|\overrightarrow{CO}^1|}{(\overrightarrow{CP})_z} = \frac{f}{(\overrightarrow{CP})_z}$$

where f is the focal length of the camera. From the definition of the camera and image plane systems, it immediately follows that the coordinates of the image of point P in the image plane are

13) $$x = \overrightarrow{CP}^1_x$$
$$y = \overrightarrow{CP}^1_y$$

However, these coordinates are in the same units as the vectors $\overrightarrow{OP}$ and $\overrightarrow{OC}$ which usually means kilometers. For purposes of specifying the rubber sheet transformation, we need to convert x and y to pixels. In order to do this, we need to know the number of pixels per unit length (such as pixels/mm) in the image plane. If K equals the number of pixels per mm in the image plane, then pixel coordinates of the image of P become

14) $$x = 10^{-6}\ K\ \overrightarrow{CP}^1_x$$
$$y = 10^{-6}\ K\ \overrightarrow{CP}^1_y$$

if $\overrightarrow{CP}^1$ is in kilometers.

Equations 5 through 14 represent an algorithm to compute $(x,y)_{in}$ from $F_C(\lambda,\phi)$. Phase 2 of MAP2 consists of the following computation for the 400 points whose latitudes and longitude were computed in Phase 1. Use Eq. 5 and Eq. 7 to compute $\overrightarrow{OP}_p$. Then use Eq. 8 to obtain $\overrightarrow{CP}_p$ and Eq. 9 to transform it to $\overrightarrow{CP}_c$. Use Eq. 12 to obtain s and then get $\overrightarrow{CP}^1$ from Eq. 10. Finally, Eq. 14 gives the desired coordinates. In order to perform this calculation, a large amount of auxiliary data is required. R_{eq} and R_{pole} of the planet in question is needed, as well as the planet-to-spacecraft vector $\overrightarrow{OC}$. The OM matrix must be known. Finally, the camera focal length and the constant K in Eq. 14 are required.

R_{eq} and R_{pole} of the inner planets and the moon are stored permanently in the program. f and K for all cameras on Mariner 9 and 10 and Viking Orbiter 1 and 2 are also built in. Their values were accurately determined in all cases with the help of pre-flight calibration data supplies by Milosh Benesh of JPL. However, the $\overrightarrow{OC}$ vector and the OM matrix vary from picture to picture in a complicated way depending on the spacecraft's orbit, time, and the orientation of the scan platform on which the cameras are mounted. The MAP2 program is relieved of much of the trouble of determining $\overrightarrow{OC}$ and OM by the existence of the SEDR (supplementary experimeter data record). One SEDR record is supplied with each picture. Contained in the SEDR is data which can be used to computer $\overrightarrow{OC}$ and the OM matrix effortlessly.

The final step in Phase 2 is to arrange the $(x,y)_{in}$ data just computed and the $(x,y)_{out}$ data from Phase 1 in the proper format. Then a general purpose rubber-sheet stretch program is fetched to produce the output picture.

VI. SPECIAL FEATURES OF THE MAP2 PROGRAM

The projection scheme described above sounds like a fairly smooth operation. If the scheme seems complicated, it is only because so many steps are involved. No single step is difficult, except the rubber-sheet stretch, but the problems encountered there are beyond the scope of this paper. However, two problems arise during Phase 2 which were ignored in Section V.

The entire discussion of Section V depends on the statement that, for Mariner and Viking cameras, the unprojected image is a perspective projection so that Equations 5 through 14 apply. However, those cameras are vidicon cameras and as such are subject to scene-dependent geometric distortion. If these distortions are ignored, mosaics cannot be made without significant mismatches at picture boundaries. The (x,y) in the preceding pictures refer to ideal, undistorted pictures. At one time pictures for mosaics had to be subjected to two rubber-sheet stretches, one to remove the geometric distortions (see Ref. 2 for details) and another for the map projection. This process increases the loss of fine scene detail caused by interpolations made during the stretch. MAP2 now has a feature by which it combines the parameters for the geometric correction and the map projection into a single composite transformation, so that a picture needs to be rubber-sheet stretched only once. The procedure it uses is simple. After Phase 2 it has

$(x,y)_{in}$ for each tiepoint. The new value is formatted into the parameters for the program fetched by Phase 3. If the user does not wish to make the effort to determine the geometric distortion correction parameters himself (using programs available at IPL based on the principles discussed in Ref. 2), a default set of distortion correction parameters is built into the program for each camera on Mariners 9 and 10, Viking Orbiters 1 and 2. This default set works very well for many pictures. In fact, most of the projections illustrated in this paper were made with the defaults. In each case the built-in transformation is based on analysis of pre-flight calibration data for the camera. For the Mariner 10 cameras in-flight pictures were also used because the geometric characteristics of the cameras changed slightly after launch.

The most serious problem not mentioned in Section V is that, while the accuracy of the $\overrightarrow{OC}$ vector in the SEDR is quite high, the accuracy of the OM matrix is insufficient for some purposes. Also, if a picture from some source other than Mariner or Viking is to be projected (e.g., earth-based moon pictures), no SEDR is available at all. All the information MAP2 needs can be entered via parameter cards, so a SEDR is not necessary. The $\overrightarrow{OC}$ vector can often be computed by hand fairly easily, but the OM matrix is again a problem. To overcome these problems, MAP2 has the ability to compute the OM matrix itself if certain additional information is given. A variety of algorithms is available. Before discussing them in turn, however, the reason for the OM matrix problem in the SEDR and the seriousness of the problem should be discussed.

Due to the design of the scan platform on which Mariner and Viking cameras are mounted, the pointing direction for each picture is only known to within some tolerance value. For Vikings 1 and 2 this uncertainty amounts to 1/4°, which corresponds to 175 pixels. The net result is that if $(x,y)_{in}$ is computed for a given (λ,ϕ) using the OM matrix from the SEDR, $(x,y)_{in}$ will be in error by up to 175 pixels. To a first approximation, this error is constant and amounts to the equivalent of a translation of the output picture. Such an error is harmless if only one picture is to be projected, or if a few pictures covering a small area are to be mosaicked by hand. But if a mosaic of a large area is desired, the second-order errors become significant even for hand mosaicking. Computer mosaics are potentially superior to hand mosaics because picture boundaries can be more easily hidden. But computer mosaic programs which assume correct projection to determine the relative offsets of the input pictures cannot tolerate even a pure translation error. Even single frames will suffer from a 1/4° error in pointing knowledge if the frame is taken from a great distance, so a large fraction of the planet appears in the field of view. In fact, for approach pictures, the 1/4° error can cause complete failure of MAP2.

Several algorithms are available in the MAP2 program to compute the OM matrix from a variety of input data. Each of them will be discussed below:

"POGASIS" method -- the OM matrix can be considered to be the product of two intermediate matrices. One of these converts from the Planetary Surface System to a standard astronomical coordinate system based on the celestial equator and the vernal equinox. This transformation depends only on the time of observation and the orientation of the target planet's orbit and rotation axis. The other matrix converts from the astronomical system to the camera system. It depends on the orientation of the camera with respect to the celestial sphere. In some early Mariner missions the first matrix above was not always accurate in the SEDR. So the MAP2 program can mimic a program called POGASIS which computes the XME matrix (as it is called) given the time of observation. This algorithm is rarely used now. Recent mission SEDR's have had very accurate XME matrices. The scan platform design problems affect the other matrix involved in the computation of the OM matrix.

QUAM algorithm (see Ref. 3) -- The Quam algorithm determines the OM matrix from the Planetary Surface coordinates of two vectors which must be known to the user. One of the vectors is the $\overrightarrow{OC}$ vector from Figure 4. The other vector is that from the planet's center to the point in the center of the image. The crux of the method is the fact that the difference between these two vectors is the vector from the spacecraft to the point in the image's center. This vector is clearly the +z-axis of the camera system. The user must supply the azimuth of north in the image also. Otherwise, only the z-axis is determined and the complete OM matrix could not be obtained.

FAR ENCOUNTER algorithm -- This algorithm can be used to project pictures which contain the entire planet (or most of it) in the field of view. Why project such pictures, since they obviously require no mosaicking? Figure 3 is an illustration of one application of this algorithm. Reference 4, a study of Venus' atmosphere based on Mariner 10 fly-by pictures, based most of its conclusions on mosaics made from small strips cut out of Mercator projections of full-planet Venus pictures. The idea was to isolate single atmospheric features and follow their time development as the atmosphere rotated between pictures. A "pseudo-Mercator" projection in which time replaced longitude as one of the variables was produced. The Far Encounter algorithm has also been used to compute the OM matrices for Viking 1 approach pictures without actually projecting the pictures. These matrices were used to determine latitudes and longitudes as functions of $(x,y)_{in}$ in the pictures. This knowledge plus solar lighting geometry from the SEDR permitted "half-Mars" pictures to be corrected for sunlight variations to produce uniformly-lit images for color processing. Figure 5 is one result of this processing.

The algorithm requires as input the $\overrightarrow{OC}$ vector in Planetary Surface coordinates, the azimuth of north measured at the subspacecraft point, and the image plane coordinates of the center of the visible disk. The details of the algorithm will be published in a forthcoming paper.

TIEPOINTS algorithm -- This is a general purpose mode which requires considerable preliminary work by the user but which can be applied to virtually any picture. If the latitude and longitude of three points of known (x,y) are available, then the OM matrix is determined. In view of the difficulty of defining and measuring the location of features of known latitude and longitude (these are often large features), it is best to measure

more than three points and determine a "best fit" OM matrix in a least squares sense. Map2 allows data for up to 20 points to be entered to compute the OM matrix. Joel Mosher of IPL has used this mode to produce a Mercator projection from lunar photos taken at ground-based observatories.

The most accurate method of all to obtain OM is to use the techniques of the science of photogrammetry to obtain accurate OM matrices and $\overrightarrow{OC}$ vectors for a large set of frames. Merton Davies of Rand Corporation supervised the required measurements and computations from Mariner 7, 9, and 10 pictures (see Refs. 5 and 6 for details). His pointing data were used for all near-encounter Mercury pictures made at IPL during Mariner 10. The pictures which make up Figure 1 were projected in this manner and subsequently computer mosaicked. Many more Mercury mosaics have been produced in this way and will be available publicly soon.

VII. SUMMARY

This paper has briefly described the primary applications of digital cartographic projections at the Image Processing Laboratory of the Jet Propulsion Laboratory. The general approach to the projection problem at IPL and some specific solutions to one of the chief subproblems, determination of the camera viewing geometry, were described. Mosaicking adjacent pictures is the most important reason for projection. However, removal of foreshortening in single pictures taken obliquely to the surface and the simplification of the determination of latitude and longitude from line and sample coordinates are also common applications of the technique.

The projection process divides into three separate phases -- locating a suitable set of points in the output (projected) picture and determining their latitude and longitude, locating the same points in the input picture, and transforming one set into the other along with the rest of the picture. The first two phases were described in some detail. The third phase is a special case of a more general problem which has been discussed elsewhere.

Finally, various methods to overcome the lack of accurate OM matrices were mentioned and some of their applications illustrated.

REFERENCES

1. Green, W. B., Jepsen, P. L., Kreznar, J. E., Ruiz, R. M., Schwartz, A. A., and Seidman, J. B., "Removal of Instrument Signature from Mariner 9 Television Images of Mars," Applied Optics, 14, 105, 1975.

2. Kreznar, J. E., "User and Programming Guide to the MM71 Geometric Calibration and Decalibration Programs," Document 900-575 (Jet Propulsion Laboratory), March 1, 1973 (internal document).

3. Quam, Lynn H., "Computer Comparison of Pictures," Stan-CS-71-219, Computer Science Department, Stanford University, May 1971.

4. Belton, M. J. S., Smith, G. R., Elliott, D. A., Klaasen, K., and Danielson, G. E., "Space-Time Relationships in the UV Markings on Venus," J. Atmos. Sci., 33, 1383 (1976).
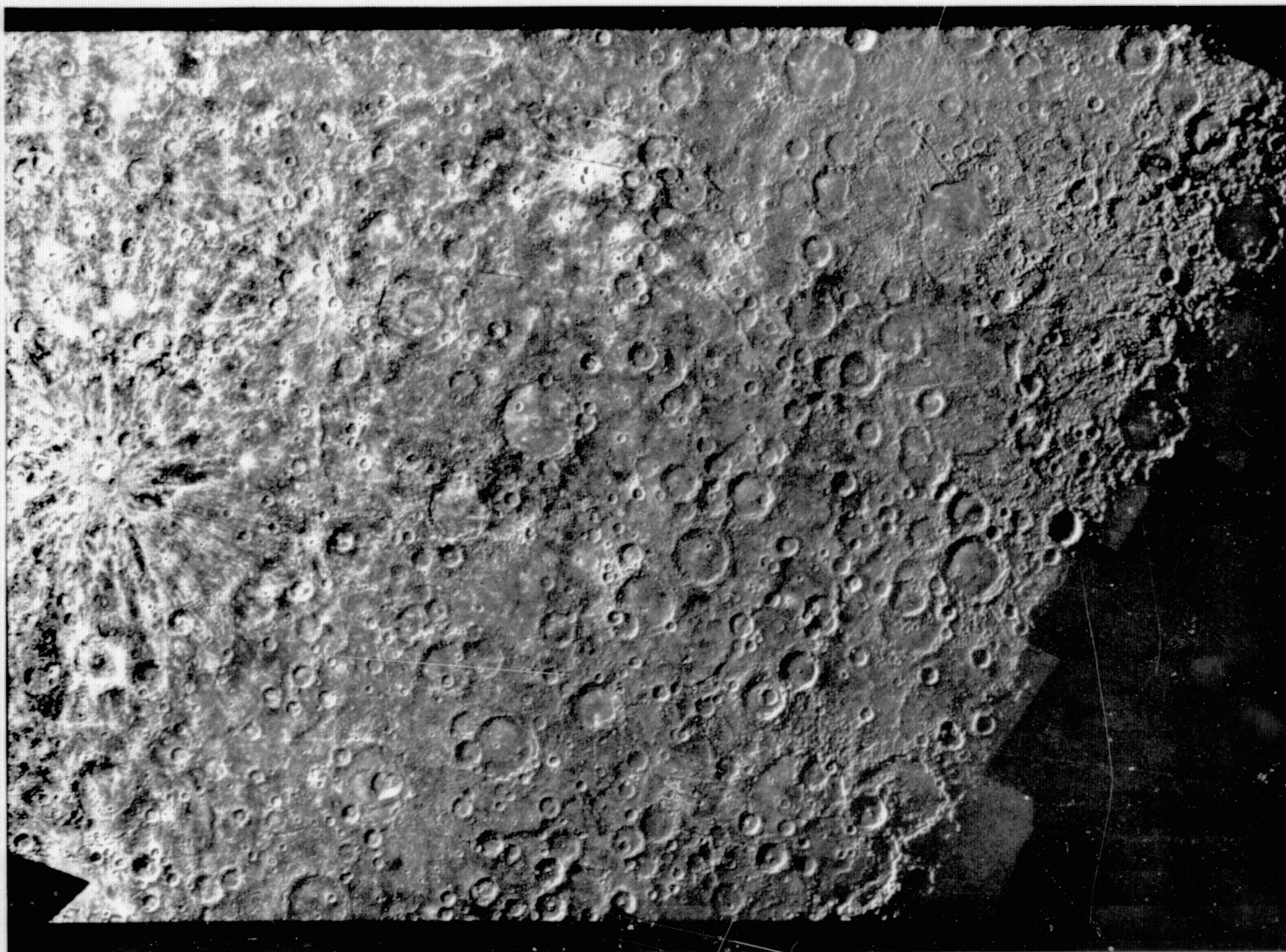
Figure 1. Quandrangle H11 of Mercury. A mosaic of 76 frames made on a Lambert Two-Standard Conformal Conic projection with standard parallel at -28° and -62° latitude, central meridian 45°W. The area included is approximately -20° to -70° latitude, 0° to 90°W longitude.
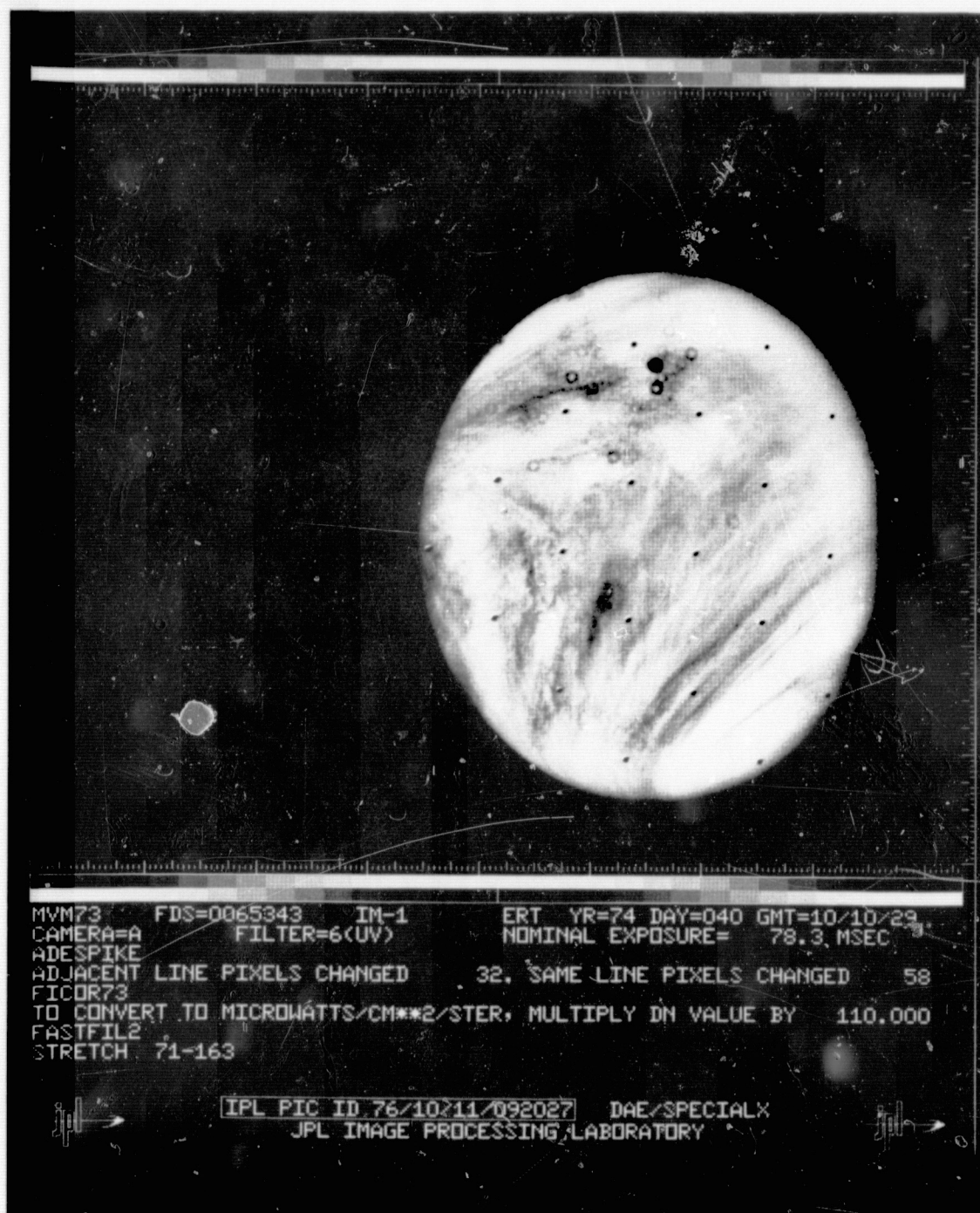
Figure 2.   A Mariner 10 picture of Venus, FDS 67543.
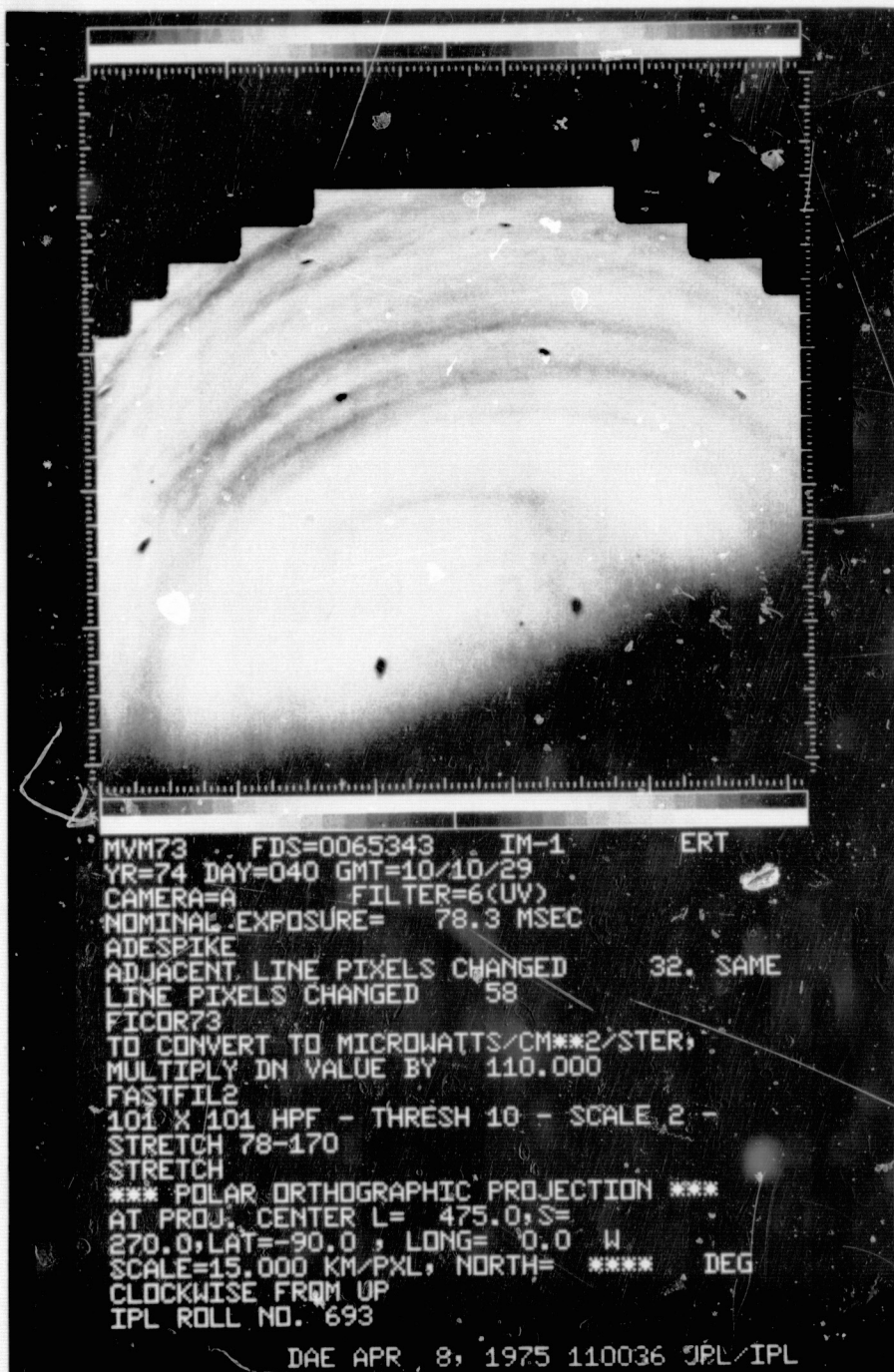
Figure 3. A Polar Stereographic projection of Figure 2, centered at South Pole.

Figure 4. Illustration of Camera Viewing Geometry

Figure 5. A color composite of 3 Viking 1 Mars pictures taken a few days before orbit insertion. The three pictures which make up the composite were the fifth in a series of violet-green-red triplets taken during approach to Mars.
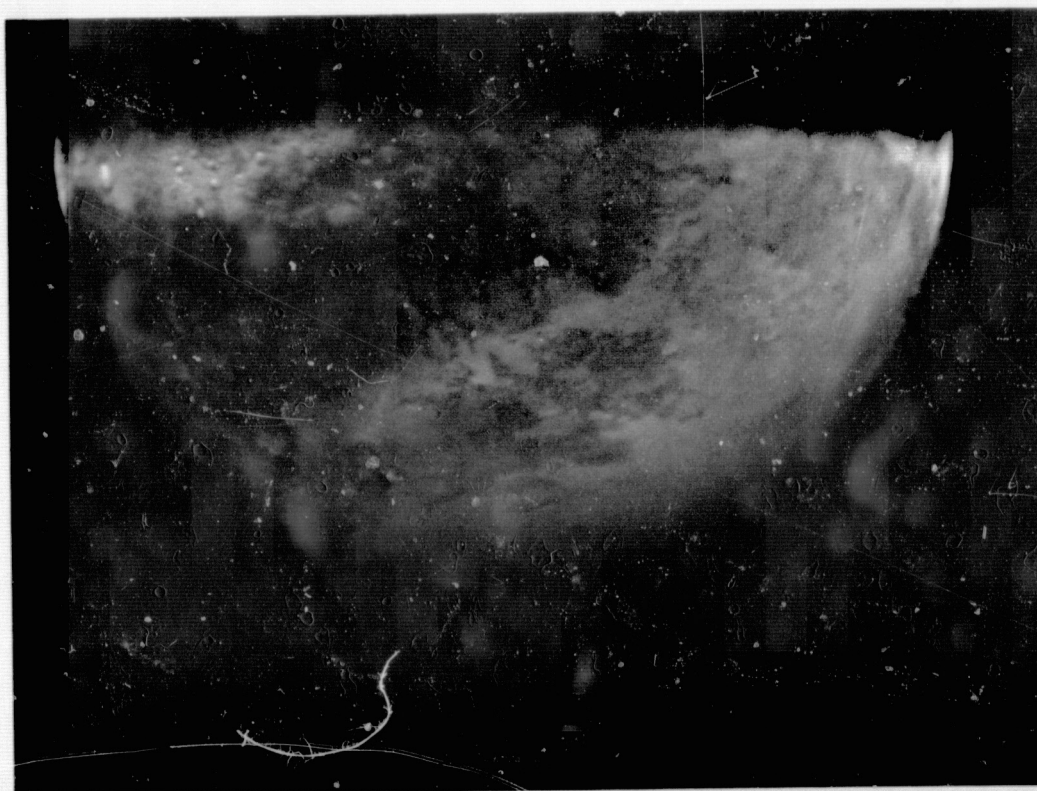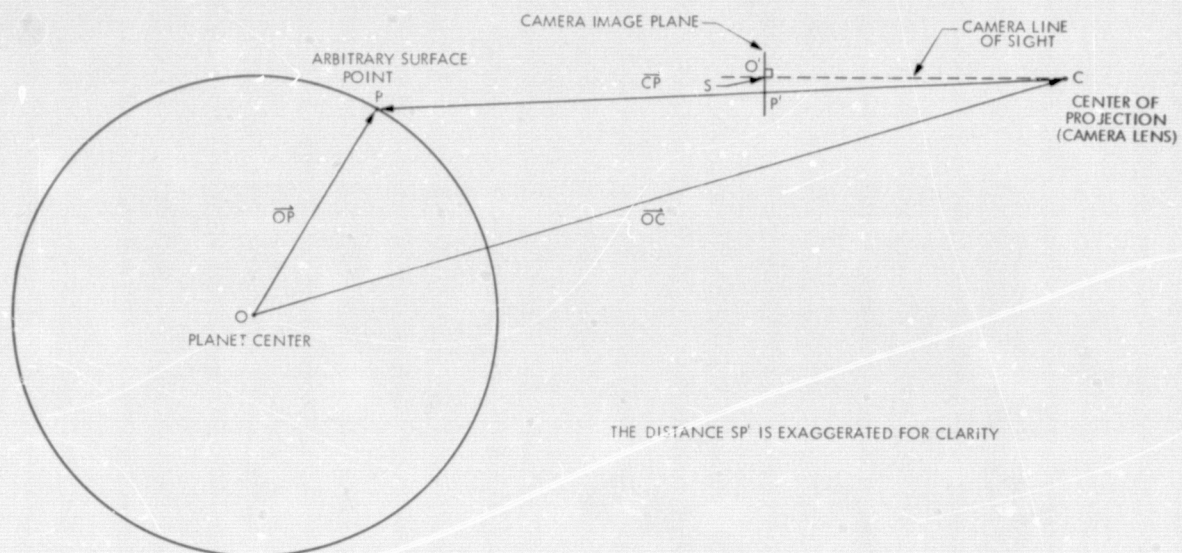
COMPUTER ANALYSIS OF CARDIOVASCULAR IMAGERY*

R. H. Selzer

Jet Propulsion Laboratory
Pasadena, California

D. H. Blankenhorn, D. W. Crawford, S. H. Brooks
and R. Barndt, Jr.

University of Southern California
School of Medicine
Los Angeles, California

ABSTRACT

A stand-alone computer image processing system has been assembled to quantify atherosclerosis from x-ray arteriograms. The processing involves scanning and digitizing the film, tracking the arterial vessel outline in the image and then computing various measures of arterial wall irregularity. The initial application of the technique was to arteriograms of the femoral artery. Two methods of edge detection (called the slope and knee method) and three types of irregularity measures (Lumen measure, Roughness measure and volume measures) have been tested by means of a post-mortem experiment which compared the computer analysis of 128 vessel segment angiograms with a direct visual assessment of the artery and with a cholesterol assay of the wall. The method has also been applied clinically to patients associated with the Los Angeles County/USC Medical Center. At the current time, the technique is being extended to the processing of coronary angiograms.

I. INTRODUCTION

In recent years, it has become known that most individuals over the age of 30 in Western Europe and the United States have some degree of atherosclerosis, the arterial disorder that results in obstruction of blood vessels from fatty plaque deposits. The cause of atherosclerosis is not completely understood but numerous associative factors have been identified (blood cholesterol levels, blood pressure, smoking) and various therapeutic hypotheses studied. However, there have been few reports of successful regression trials in which vessel improvement could be identified with a specific type of therapy. The problem appears to result from the lack of a practical method to detect vascular change. Traditional epidemiological methods require large number of patients and long trials and while angiography has been used to demonstrate the existence of atherosclerosis, attempts at visual quantitation have not been sufficiently accurate to detect short-term change. In addition, most angiographic studies have been concerned with advanced or near-obstructive disease which may in fact be the stage of the disease least responsive to therapeutic procedures.

It is with this background that JPL and the USC School of Medicine undertook the development and testing of a computer method to quantify atherosclerosis from angiographic film. This was accomplished initially by means of a grant from the National Heart and Lung Institute to establish at USC a Specialized Center of Research (SCOR) for Atherosclerosis with USC conducting a clinical and experimental program and JPL developing the needed computer-based diagnostic instrumentation. In addition to the above, support has been received this year from the NASA Office of Life Science to establish a Medical Image Analysis Facility at JPL to study the effects of stress in persons with coronary atherosclerosis. The goals of this program are closely related to those of the SCOR Program and involve both USC and the Ames Research Center in a collaborative program.

The specific therapeutic procedures selected for evaluation in the USC SCOR study were weight loss and exercise. The individuals recruited for the program were men ages 40-49 who had experienced a prior myocardial infarction. The objective of the computer diagnosis was to detect atherosclerosis changes in these subjects from angiograms taken in one or two year intervals.

For reasons explained below, the initial computer analysis was applied to femoral angiograms. The computer assessment of atherosclerosis is based on a measurement of the roughness or irregularity of the lumen shadow in the angiogram. An interactive, stand-alone computer image processing system has been assembled and programmed to perform this task. The technique has been tested in a post-mortem experiment comparing the computer analysis with a visual assessment technique and with a cholesterol assay of the vessel wall. The method has been applied to films from subjects in the regression study described above and also to patients from the Cardiac Lipid Center at Los Angeles County Hospital.

In the remainder of this paper, a description is given of the image processing system and the computing techniques. The experimental and clinical aspects of the study are briefly mentioned but the specific medical results of the program are beyond the scope of this discussion (Ref. 1-6).

II. IMAGE PROCESSING SYSTEM

The image processing system assembled to perform the atherosclerosis assessment is shown in Figure 1. The computer and peripheral components are described in the Appendix. A Dicomed Corp. D57 Film Digitizer is used to convert the angiographic image to a number array. As shown in Figure 2, transmitted light from the film reaches the photocathode surface of an electronic image dissector which emits an electron cloud whose intensity is proportional at each point to the incident illumination. By deflecting this electron cloud past a small aperture, the image is scanned and digitized. The primary difference between this device and the more common television systems using vidicon tubes is that the image dissector uses a static deflection system that allows any image point to be directly addressed and then sampled for either long or short periods as needed to achieve a given signal-to-noise level.

The D57 can generate a sampling raster of 2048 lines by 2048 points per line to cover a film area of 54 mm by 54 mm. The minimum sample spacing is thus 54/2048 = .026 mm with selectable steps of .052, .104 and .208 mm available. Four sampling speeds ranging from 20 microseconds per sample to 1.28 milliseconds per point are available. A 54 mm by 54 mm area can be converted to a 256 x 256 digital array using the 20 microsecond sample rate in 2.3 seconds while a 2048 x 2048 array covering the same area but using the 1.28 millisecond sampling time requires 1.5 hours.

Output from the film digitizer is fed directly to the computer core memory by direct memory access and transferred to disk storage at the completion of each line. At the same time, each digitized line is also transferred to the memory of the Ramtek 100B display system, whose output drives a TV monitor. Thus, the digital image is displayed to an operator as the scanning proceeds. The Ramtek MOS memory holds a 640 sample by 512 line picture with each picture element represented by 4 bits (16 grey levels). A fifth bit provides a graphics overlay needed for various interactive aspects of the processing. The use of the Ramtek system as well as the Science Accessories sonic digitizer in an interactive mode will be described below.

Hard copy pictorial output is obtained from an Image Information, Inc. recorder which uses an intensity modulated laser to expose dry silver paper. The unit is capable of recording 250 picture lines per inch of paper and generates 16 shades of grey. A 1024 line picture requires approximately 70 seconds to be recorded and developed.

III. ASSESSMENT METHOD

The processing of an individual arterial segment can be viewed as a four step operation involving the following parts: (1) angiography, (2) film scanning and digitizing, (3) edge tracking, and (4) irregularity measurement. Each of these steps are described below.

1. Angiography

At the outset, the decision was made to develop the method first to assess disease in the femoral artery and then extend the technique if necessary to the coronary arteries. Although both vessels are prominent sites for atherosclerosis, the degree to which the disease is related in the two arteries was largely unknown when this study began. Since coronary arteriography does have a higher patient risk, and since the image processing task for the femoral artery was less complex due to the fact that the femoral is essentially stationary and surrounded by radiographically near-homogeneous tissue, the femoral was chosen as the starting point. Since the program began in 1971, evidence has been accumulated that coronary and femoral atherosclerosis is not closely enough correlated to allow prediction of coronary disease or change from the femoral data alone. As a result, extension of the technique to coronary arteriograms has begun. The description below covers the development of the femoral processing method while the newer coronary processing technique is described in Section V.

An example of a femoral arteriogram is shown in Figure 3. In a given examination, approximately 20 films are taken over a ten second period following contrast injection and a particular film is selected for computer processing only after determination that the vessel is completely filled with contrast material and that no evidence of flow separation or other artifacts exist.

These examinations are repeated at approximately yearly intervals with care taken to maintain constant position and contrast. An example of the degree of constancy which can be achieved is shown in Figure 4.

2. Film Scanning and Digitizing

The maximum scan area of the Dicomed film digitizer is 54 mm x 54 mm so the full length of the femoral artery image is covered in three to five steps. The position of each 5 cm segment is indicated by a clear plastic overlay taped to the film by the cardiologist. These overlays have positional markings related to the digitizing equipment that the computer operator utilizes during setup. Each femoral segment covering 27 mm x 54 mm of film area is converted to a 512 x 1024 array of eight-bit (0-255) numbers approximately linearly proportional to film density.

3. Edge Tracking

The first processing step applied to the digitized femoral picture is edge tracking. A fundamental characteristic of an x-ray image is that it is extremely noisy compared to most optical images. As a result boundary detection algorithms must either (a) be applied to images that have been pre-smoothed, (b) include a large enough portion of the picture in the determination of each boundary point so that an effect of data smoothing is obtained, or (c) smooth the detected boundary points. We have utilized (b) and (c) to find the vessel edges.

In order to smooth over enough points and at the same time prevent the processing from becoming too lengthy or complex, we perform a linear rather than two-dimensional search for edges. Since a 5 cm segment of the femoral artery is usually nearly straight, the vessel image is oriented vertically when digitized and the search for edges made along horizontal scan lines which are essentially perpendicular to the vessel. This orientation is shown in Figures 5a and 5b.

Once the direction of search is determined, a decision must be made as to which point along the search path in the image corresponds to the real vessel edge. In an optical imaging system, an opaque object creates a shadow decomposable into a reasonably predictable numbra and penumbra. Given sufficient knowledge of the imaging system, a correspondence between image points and object points can be established. For x-ray images, this is a difficult task for a known object such as a normal vessel and is in fact impossible for an unknown object since the image results from differential absorption of radiation and the concepts of numbra and penumbra are not applicable. In order to establish the exact correspondence between image and object points for an x-ray system, it is necessary to know not only the imaging geometry and recording media response, but also the following:

(1) The size, shape and intensity distribution of the x-ray source.

(2) The spectral energy composition of the x-ray beam.

(3) The size, shape and chemical composition of both the object to be mapped and all nearby objects that scatter x-rays.

For clinical applications items (1) and (2) can be measured but with difficulty. Item (3), for an irregular vessel, is by definition unknown. As a result, the real vessel edge cannot be found and it is necessary to find some other image feature that reflects atherosclerotic change. Methods for detecting two such features, called in this paper the slope method and the knee method, have been developed and tested and will be described below.

Slope Method. The edges of the vessel image, using the slope method, are taken as those points along a horizontal scan line where the gradient of the intensity values is maximum. Consider Figure 6a which shows a typical 512 point intensity profile through the vessel in the orientation shown in Figure 5b. The image of a small branch also appears on the left-hand side of Figure 6a. The gradient at any given point along the line is calculated as the derivative of a second degree least-square polynomial fit through the eleven nearest intensity values on each side of the given point. Figure 6b shows a plot of this derivative obtained by repeating the above calculation at each point along the line. This "moving" calculation is shown diagramatically in Figures 5c and 5d. The left vessel "edge" is selected as the coordinates of maximum positive value of the derivative and the right edge as the maximum negative value.

Figure 6b shows the derivative calculated for almost the entire 512 point line. In practice, the search for the maximum gradient is constrained

on each side of the vessel to 31 points with the location of the previous edge (one line up) determining the center of the search area.

The actual computation of the derivative of the second degree least square polynomial is quite simple. If $X_1, X_2 \ldots$ are the intensity values along the scan line, S_n the derivative at the nth point of the polynomial thru the nearest $\pm K$ points ($K = 11$ as described above), then

$$S_n = C_k \sum_{k=-K}^{K} k \, X_{n+k}$$

where C_k = constant dependent on K. Further, if a quantity P_n is defined

$$P_n = C_k \sum_{k=-K}^{K} X_{n+k}$$

Then, a simple recursive computation is possible,

$$S_{n+1} = S_n + P_{n+1}$$

and each new derivative is obtained from the old with two addition and one subtraction operations.

An application of the slope method of edge detection is shown in Figure 7. Note that the intersection of the main vessel with a branch or with a crossing vessel is not marked with edges. These areas are deliberately skipped or erased interactively by the operator since the edge location becomes ambiguous. These portions of the edge are excluded from subsequent computation of vessel irregularity measures.

Knee Method. The knee of the vessel profile along a horizontal scan line is loosely defined as the point of intersection of the vessel image with the background image. For the reasons discussed earlier, these intersection points cannot be precisely defined, particularly when the vessel has an irregular cross-section, so an empirical approach was taken. One method that seemed to work well was to select the knee location as the coordinate of the maximum rate of gradient change, i.e. the second derivative. The computation was accomplished by applying the slope method twice, first to the original profile data and then to the resulting first derivative output. The result of this type of computation is shown in Figures 8 and 9. A 21 point slope calculation (instead of 23) was used in this calculation so the effective smoothing length of a double application of the slope method is 41 points.

The slope method has an intuitive appeal because it places the vessel edges at about the same place a human observer would select visually. The knee method, on the other hand, more accurately reflects the actual edges of vessel image, at least

for the case of a near-normal artery. The knee method edges are less sensitive to noise and contrast variations because of the additional data smoothing involved but as a result require longer to compute (200 seconds minimum for the knee method to track 1024 lines vs. 70 seconds minimum for the slope method). However, as will be shown later, the irregularity measurements were not significantly different for the two edge tracking methods.

4. Vessel Irregularity Measurement

Once the vessel edges have been located, the next processing step involves quantitation of the vessel irregularity or roughness. Our experience has been that no single measurement is adequate but rather a series of measurements are needed. For example, an algorithm that detects and accurately quantifies a large fibrous lesion will probably not recognize early diffuse fatty streaks.

The measurements can be broadly categorized into those that are primarily derived from edge information and those primarily derived from optical density variations within the lumen image. These latter measures are referred to as "volume" measures. Edge-type irregularity measures have the disadvantage of being rotation-dependent but are generally easier to compute and less sensitive to image contrast variations than are volume measures. Volume measures are theoretically rotation-invariant, but as described later, digitizing problems have made it difficult to fully utilize the volume measures. The edge-dependent measures described below include Roughness, Lumen 90, Tapered Lumen 90 and various width measures. The volume measures follow in the discussion.

Roughness Measure. The angiogram provides information only about the current vessel lumen - plaque deposits as well as the original arterial wall are not visible in the film since their x-ray absorption properties are nearly the same as that of the surrounding tissue. As a result, inferences about the normal vessel wall must be drawn entirely from the information about the existing diseased vessel edges. Several methods for deriving such a "normal" reference have been developed. Consider the edge representation in Figure 10a. The overall vessel curvature as well as the localized edge variations due to disease must be determined. One method for estimating the latter is to smooth the edges over a sufficiently large number of points so that all edge deflections except the overall vessel curvature are removed. An example of this type, called a long filter, is shown in Figure 10b. The coordinate of the edge point on the right, along the dashed line, is obtained by averaging the coordinates of the edge points enclosed within the box. The filtering is repeated for each edge point. An 11 point filter was specified for illustrative convenience but in actual practice, smoothing filters of this type cover several hundred edge points.

Repeating the filtering operation on the original data, but with a short filter, as shown in Figure 10c, removes the effects of noise. A measure of irregularity called the Roughness measure is defined as the root-mean-square difference between the long and short filter, as shown in Figure 10d. For the illustration, the Roughness

would be labeled R(5,11). The equivalent measure calculated for an actual femoral image, intended to compare detected edges less noise with vessel curvature is R(3,321). All subsequent discussion of filter sizes will refer to actual femoral processing rather than to the illustration.

This same method, using different filter sizes, can be used to tune the measurement to a particular lesion size. For example, an R(3,21) filter will discriminate against some edge noise but will not remove the effects of small plaques. An R(91,321) filter will tend to be more sensitive to large lesions and will ignore small plaques. As discussed later, a post-mortem experiment was conducted to determine the optimum filters.

Lumen measure 90. The Lumen measure 90, L90, which uses both edges to generate a computer estimate of a normal or pre-disease vessel lumen is calculated as follows: The left and right edges for each line are averaged to obtain a sequence of vessel midpoints. These midpoints are then smoothed with a 401 point filter to obtain a vessel midline (see Figure 11). This midline is then translated without rotation to the left until it encloses 90% of the left edge points. The translation of the midline is repeated to the right to establish a curved cylinder that encloses 90% of the edges on each side. The root-mean-square difference between the original edges and this computer generated lumen is the L90 measurement.

Tapered Lumen 90 measure. The Tapered Lumen 90 measure, TL90, is derived exactly like L90 except a tapered cylinder is derived. The amount of taper is determined from the slope of a regression line through a plot of vessel widths as a function of distance along the vessel. The computer midline, as described above, is made to tilt by half the taper value before translation on each side. The remainder of the computation is the same as for L90.

L90 and TL90 illustrate the need for multiple irregularity measures. A normal, tapered vessel will be correctly assessed by TL90 but incorrectly by L90. The opposite situation occurs if a vessel has progressively greater disease from top to bottom. The latter condition does not occur very frequently and our experience has been that TL90 is overall a more useful measure than L90.

Width measure. Minimum, maximum and average vessel width are measured in order to compute maximum stenosis, detect hemmorhage and to detect generalized uniform progression or regression.

Volume measures. Computing the mean intensity of the vessel image along a given horizontal scan line produces an approximate measure of cross-sectional vessel area or incremental volume. By computing the variability of this volume along the vessel length, a new set of irregularity measurements, potentially invariant under vessel rotation, can be defined.

By computing the mean line intensity from the left computer lumen (as calculated for the L90 measure) to the right, the estimated volume per line is relatively independent of the edge tracking, at least to the extent that minor localized

changes in the detected edges, perhaps from noise, will have little effect on the computer lumen and thus little effect on the mean line intensity. The root-mean-square variation of these lumen-to-lumen mean line intensity values about the average for the entire vessel segment is defined as the Outer Density Average (ODA) measure. The Outer Density Sum (ODS) measure is defined as ODA times computer lumen width. Two additional volume measures were defined as above, but the mean line intensities were computed between the detected edges instead of from the computer lumen. These measures were called Edge Density Average (EDA) and Edge Density Sum (EDS).

IV. SYSTEM TEST AND VALIDATION

In the previous section, two methods of edge tracking and a number of approaches to the problem of quantifying vessel irregularity were described. In this section, some of the methods for testing the algorithms to determine the best approach amongst the various options is discussed.

The primary experiment conducted to test the processing methods was a post-mortem study of 128 femoral vessel segments from 21 cadavers. Femoral angiograms were made under simulated clinical conditions using a pressurized radiopaque silicon rubber casting technique. The vessels were excised, opened and photographed in color alongside the cast and the arterial wall was then analyzed for cholesterol content. Four graders on two occasions sequenced the photographs in increasing order of disease based on the International Atherosclerosis Grading Scheme. Correlation between the two grading sessions was .984. The visual grading and the cholesterol assay were then compared with the computer analysis of the angiograms. The computer applied both the slope and the knee edge methods to the images and tested the following irregularity measures: L90, TL90, average width, maximum stenosis (average width minus minimum width divided by average width), volume measures, and a matrix of 45 roughness measures where one filter varied from 3 to 129 points while the other filter varied from 21 to 321 points.

For each edge method, all of the irregularity measurements were allowed to compete in a stepwise regression analysis first against the visual grading and then against the cholesterol assay. This experiment has been reported in detail in previous publications (Refs. 7 and 8) but the portion of the results pertinent to this discussion can be summarized as follows:

1. The highest simple correlation of a computer measure against visual grading was .76 for TL90 with slope method edges. Other measures including R(33,321), R(97,321), R(81,97), L90, maximum stenosis and the volume measures were also correlated at levels from .52 to .72 at a significance level of .001.

2. The highest simple correlation of a computer measure against cholesterol content was .75 for R(321,97) with R(321,33), R(321,3), L90, TL90, ODA and ODS all having correlation values above .70 at a significance level of .001.

3. An optimum linear combination of slope method computer measures involving TL90, R(81,97) and R(3,21) produced a multiple correlation level of .85 against visual grading. Comparable result for computer vs. cholesterol was .78.

4. The slope and knee edge detection method results were not significantly different in the experiment.

5. Volume measurements did not add significantly to the multiple regression correlation.

From the above results, two new computer measures, derived from the stepwise regression, have been added to the processing sequence. These are Computer Estimate of Atherosclerosis, CEA and Computer Estimate of Cholesterol, CEC. Each of these is a linear combination of the most significant individual irregularity measurements.

An example of the standard output of a processed clinical film using the irregularity measures described above is shown in Figure 12. The irregularity measures along the sides are computed over one centimeter sections of the image, as delimited by the horizontal lines.

To further test the system, various replication experiments were conducted as described below.

Scanner and operator variation. A single segment on one angiogram was processed eleven times over a three month period using three different operators. This test was designed to measure the effects of long term electronic changes in the Dicomed film digitizer and operator variability during film digitizing or during the interactive edge tracking. Tables 1 and 2 show the resulting data for selected measures using the slope and knee edges, respectively. The coefficient of variation of the single measures for the slope method runs from 2.1 to 7.0% while that for the knee method runs from 2.7 to 10.2%. The multiple prediction measurements derived from the visual grading and cholesterol analysis in the post-mortem experiment run from 2.0 to 3.9% for the slope method and 4.2 to 4.9% for the knee method.

Exposure variation. To test the effect of film contrast on the algorithms, two phantom vessels were fabricated, filled with Renografin at five levels of concentration varying from 10% to 50% (4.8% iodine to 24.2% iodine) and radiographed at 75, 80 and 85 KVP. One phantom consisted of a negative cast of a moderately diseased vessel filled with Renografin. The second phantom consisted of a smooth cylindrical channel to simulate a normal vessel.

The computed values of L90 and R(97,321) using slope method edges are shown in Table 3 for the two phantoms. Note the general decrease in L90 values as a function of increased Renografin concentration, particularly for the smooth tube phantom. The increased vessel image-to-background contrast ratio associated with greater concentration of Renografin results in an effective decrease in image noise. Since L90 measures for a smooth tube should be zero in the absence of noise, a decrease in their values

as Renografin increases is expected. The moderate disease phantom also shows this effect but to a lesser extent since the zero-noise irregularity is not zero. The standard deviation of the measurements for each phantom are comparable, but the coefficient of variation is lower for the moderate disease phantom because of the larger average values of irregularity.

The multiple measures, CEA and CEC, which predict visual atherosclerosis and cholesterol respectively are shown in Table 4 for the same Renografin and KVP values as above. Note that the percent variation for these combined measures is in the order of 3% for the moderate disease phantom.

Vessel image contrast, defined as the ratio of the optical density difference between the vessel center and the background, divided by the background optical density, was measured and found to vary for the smooth tube films from a value of .378 (at 10% Renografin, 85 KVP) to .651 (at 50% Renografin, 80 KVP). Comparable values for the moderate disease phantom were .296 to .672 (same concentrations and KVP values).

Angiographic replication. The variation in the computer analysis of the human femoral artery due to such factors as the degree of dye mixing and possible pulsatile effects in the artery are relatively difficult to measure. The ideal test would be to repeat the entire examination in an interval too short for measurable change in the disease - for example, a day or one week later. However, considerations of patient risk prevent this from being done. A partial solution to this problem is to process multiple films from the same examination. In this case, the time interval between frames is in the order of seconds or less and while not all possible variations can be tested, data can be obtained on such effects as dye mixing. One experiment conducted along this line involved the processing of film pairs from six patients. The pairs were selected to exclude obvious mixing defects, as is normally done before processing clinical films. The results of processing for these film pairs, using both slope and knee method edges is shown in Table 5. The values shown represent the standard deviation of the six differences values divided by the mean of the 12 computed measures, expressed as a percent. Note that the irregularity measures based on knee method edges, with coefficient of variation ranging from 2.2 to 6.1 percent, performed better than the slope method measures whose variability ran from 5.7 to 8.8%. It should also be noted that the coefficient of variation of image contrast between pairs of films was 4.8%.

Digitizer variability. The response of the film digitizer over various parts of the scanning area is variable due to non-uniformities in both the illumination source and in the image dissector. Further, there is a fatigue effect in the photocathode phosphor that causes the image dissector output to continually decrease over a period of time as long as the tube is illuminated. The amount of this decrease varies over the surface of the field of view and also as a function of the intensity of the illumination. No practical method has been found to calibrate the combined inherent spatial variation and phosphor fatigue

to the extent necessary to obtain reliable volume measurements. However, techniques for standardization of the digitizing operation in terms of the illumination levels and the length of time the image dissector is exposed to the film have helped to minimize error to the edge-dependent algorithms but it is clear that some contrast-dependent error exists.

Discussion of system test. A conclusion that can be drawn from the system test and validation procedure is that the knee edge method does not appear to be a major and significant improvement over the slope edge method and since it is considerably slower to compute, the slope method has been selected for the initial clinical phase of the program.

It also appears at this time that in light of the regression analysis of the 128 segment postmortem study and because of the difficulty in controlling or calibrating the intensity response of the image dissector digitizer, the volume-type measurements do not substantially improve the analysis. However, while edge measurements may provide a suitable sampling of disease in a group study such as that described above, on an individual basis, a rotation invariant measure would be preferable and work will continue on the volume-type measures when the digitizing problems are solved.

An exact overall figure of precision or accuracy of the analysis has not been derived but the partial test results described above suggest that a figure of less than 10% is not unreasonable, particularly after some identified sources of error are eliminated. One example of this sort involves the current use of a stationary collimating grid to reduce x-ray scatter. The faintly superimposed image of this device on the film is not noticed by the human observer but is detected by the computer and causes obvious low-level noise to appear in the detected edges. A spinning grid which eliminates this problem has been tested and found to considerably decrease the edge detection noise. An equipment change of this sort could not be made in the middle of the on-going clinical trial because the yearly films must be as nearly identical as possible. However, the spinning grid has been used in autopsy studies and will be used in any future trial.

It also seems clear that much better results could be obtained in the future if the non-uniformities in the film digitizing process could be eliminated. For example, a stable, spatially uniform digitizer, such as a drum-type film scanner would make it possible to precisely measure and correct for film-to-film contrast differences that affect both edge-dependent and intensity-dependent measures. In addition, such a system could also be calibrated to allow computation of real vessel volume.

V. CORONARY VESSEL TRACKING

A primary objective of the study at the current time is to extend the femoral techniques to coronary angiograms. The coronary tracking program that has been developed differs from the femoral programs in several important ways. As previously

described, the femoral artery is nearly straight and the arteriogram can be vertically oriented for digitizing. The coronary arteries on the other hand usually curve and in addition, bifurcate into smaller branches. As a result of these differences, a program was developed to track arbitrarily oriented curved vessels. This processing approach involves a higher degree of operator interaction than that for the femoral. For example, an operator must indicate on the digitized image which vessel in the frame to track and must also indicate the approximate path of the vessel. This is accomplished by pointing to the middle of the vessel at a series of points with the sonic digitizer, as shown in Figure 13. When this operation is complete, the computer generates a smooth continuous curve through these points (see Figure 14) and then searches along a path orthogonal to this curve for the edges of the artery. The tracked vessel for the path in Figure 14 is shown plotted in Figure 15. The maximum intensity gradient (slope method) was used to identify the edges. Finally, the computer generated lumen is found and shown plotted in Figure 16.

VI. SUMMARY AND FUTURE PLANS

A computer image processing technique to quantify atherosclerotic changes in the femoral artery of human subjects has been developed. The method has been tested in a post-mortem experiment and found to be highly correlated with directly visual observation of the excised artery and with the cholesterol content in the vessel wall. The technique has been applied to patients in a regression study and to patients from a Cardiac Lipid Center. These results are currently being analyzed and the technique is being extended to the analysis of coronary angiograms.

In addition to the above, JPL, USC and the Ames Research Center are just beginning a cooperative study of the effects of stress on persons with cardiovascular disease. The approach taken is to develop a computer/x-ray technique to estimate the maximum flow capability in coronary arteries containing diffuse complex disease. This program will involve the application of image processing methods to both coronary artery and ventricular angiograms and will utilize both analog and digital processing techniques.

APPENDIX

SCOR Image Processing System Computer Configuration

The PDP 11/45 computer used for the SCOR Image Processing System has 64K bytes of DEC core memory plus 64K bytes of add-on core memory from Standard Memories, Inc. The CPU also contains a DEC FP11B floating point processor and three DR11B Direct Memory Access units have been included to communicate with the Ramtek monitor, the Dicomed digitizer and the Image Information, Inc. laser recorder. The system utilizes a DEC RK11/RK05 disk platter (2.5 Mbyte capacity) plus two DIVA, Inc. DS18 disk units (each with 10 Mbyte capacity). Two DEC nine track magnetic tapes and a DEC seven track magnetic tape complete the system.

REFERENCES

1. Blankenhorn, D.H.: Studies of regression/progression of atherosclerosis in man. To be included in the Proceedings of the International Workshop - Conference on Atherosclerosis. Edited by G.W. Manning and M. Daria Haust. Plenum Publishing Corp., In Press.

2. Sanmarco, M.E. and Blankenhorn, D.H.: Computer analysis of femoral angiograms for evaluation of atherosclerosis in post-infarct males. Clinical correlates. Proceedings Conference on Cardiovascular Imaging and Image Processing: Ultrasound, Angiography and Isotopes. Stanford University, 1975.

3. Blankenhorn, D. H.: Femoral angiography to evaluate hyperlipidemia therapy. To be included in Lipid Pharmacology, Volume 2. Edited by R. Paoletti and C.J. Glueck. Academic Press, Inc.

4. Barndt, Blankenhorn, and Crawford: Prevalence of asymptomatic femoral artery atheromas in hyperlipoproteinemic patients. ATHERO. 20: 253, 1974.

5. Sanmarco, Hanashiro, Selvester, and Blankenhorn: Clinical arteriographic correlations in asymptomatic men post infarction. Included in Coronary Artery Medicine and Surgery. Edited by J.C. Norman. Appleton-Century-Crofts, publisher, 1975.

6. Barndt, R., Blankenhorn, D.H., Crawford, D.W., and Brooks, S.: Femoral atherosclerosis regression/progression in treated hyperlipoproteinemic patients. Annals of Internal Medicine, accepted for publication.

7. Crawford, D.W., Brooks, S.H., Selzer, R.H., Barndt, R., and Beckenbach, E.S.: Computer densitometry for angiographic assessment of arterial cholesterol content and gross pathology in human atherosclerosis. J. Laboratory & Clinical Medicine, accepted for publication.

8. Brooks, S.H., Crawford, D.W., Selzer, R.H., Blankenhorn, D.H., and Barndt, R., Jr.: "Discrimination of human arterial pathology by computer processing of angiograms," (in preparation).

SCOR IMAGE PROCESSING SYSTEM
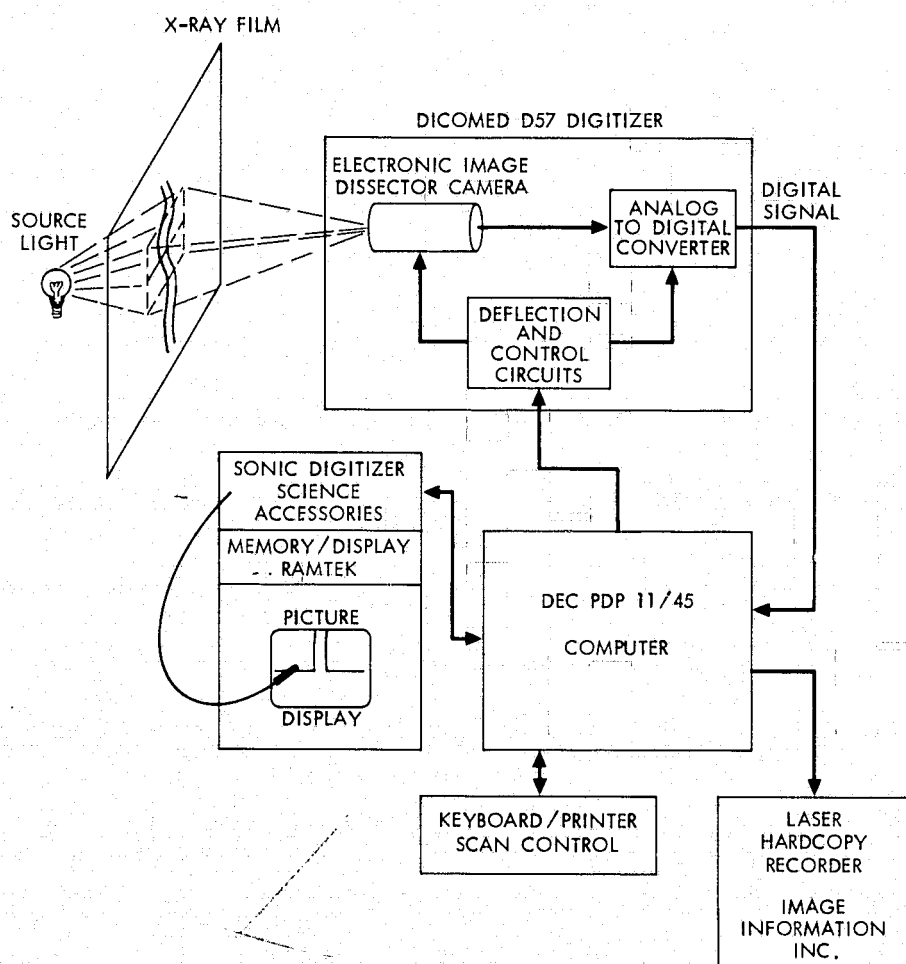
Figure 1. SCOR image processing system.
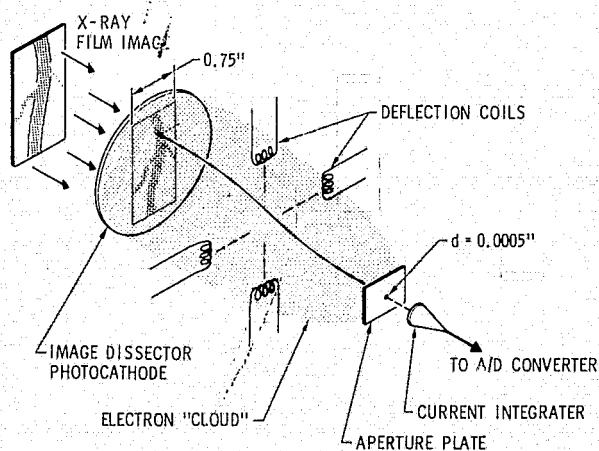
Figure 2. Electronic image dissector.

Figure 3. Femoral arteriogram.
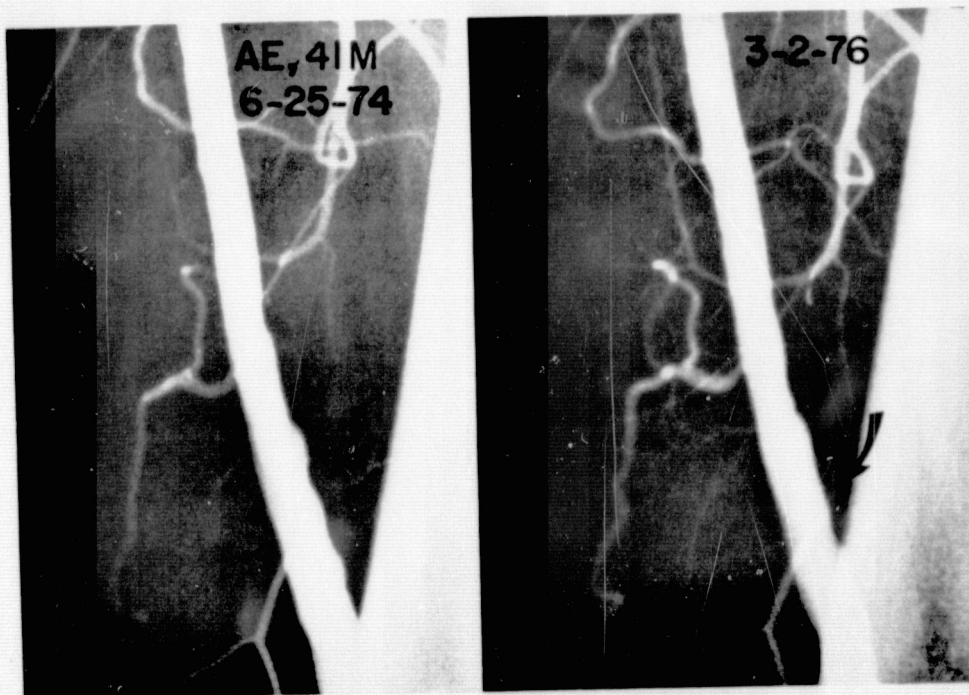
Figure 4. Femoral arteriograms taken approximately two years apart showing constancy of contrast and orientation.

(a)

(b)

(c)

(d)
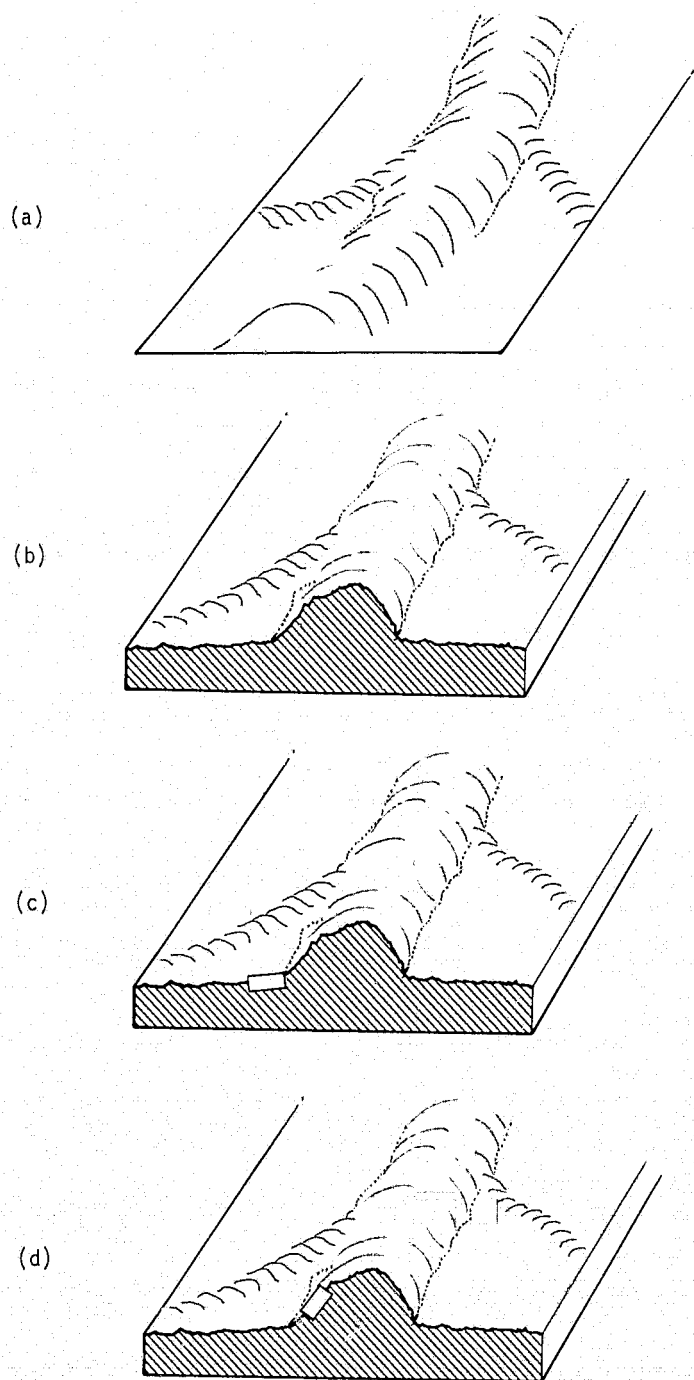
Figure 5. (a) Three dimensional drawing of angiographic
image.

(b) Cross-section view along one scan line.

(c,d) Representation of search direction for edge
tracking programs.

(a)

(b)
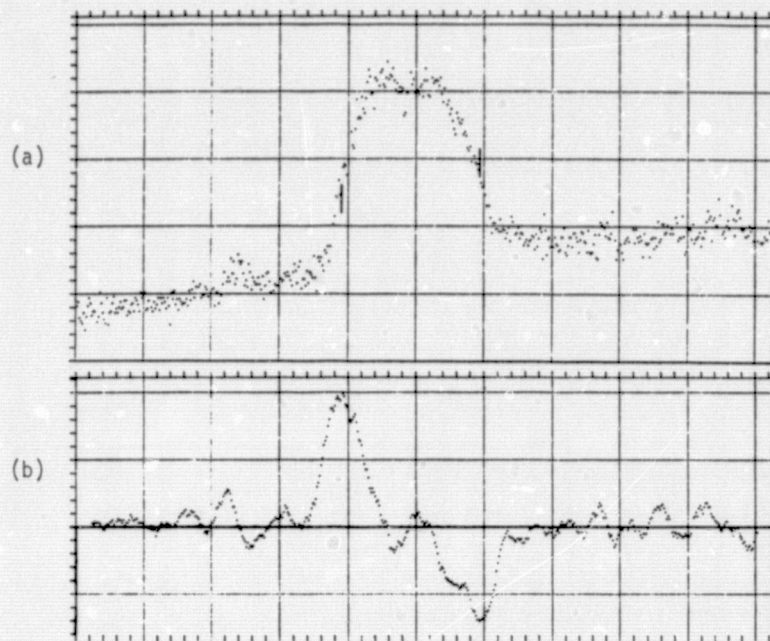
Figure 6. (a) Intensity profile along horizontal scan line with slope method edges marked.

(b) Gradient plot for (a) using 23 point second-degree least-square polynomial.

(a)

(b)
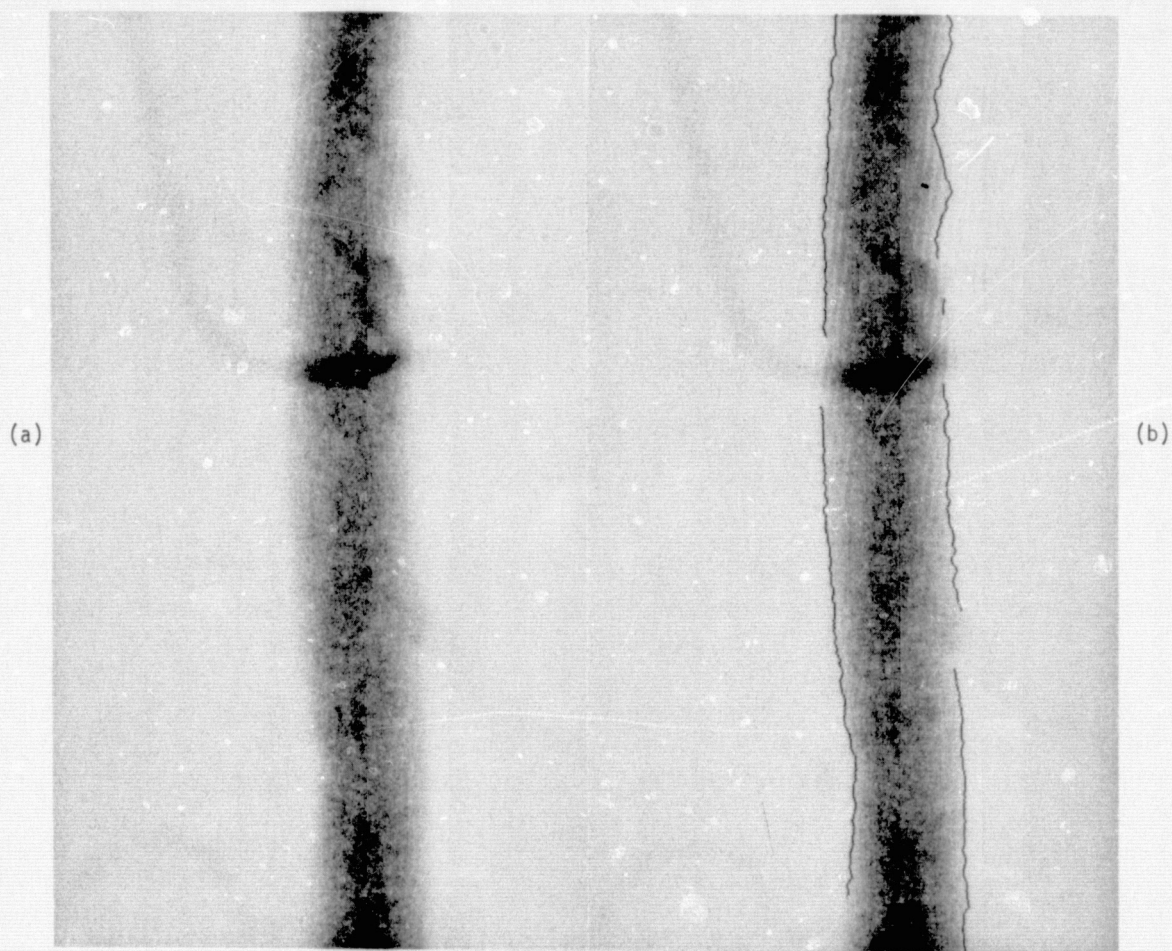
Figure 7. (a) Femoral segment before edge detection.

(b) Femoral segment showing slope method edge detection.

(a)

(b)
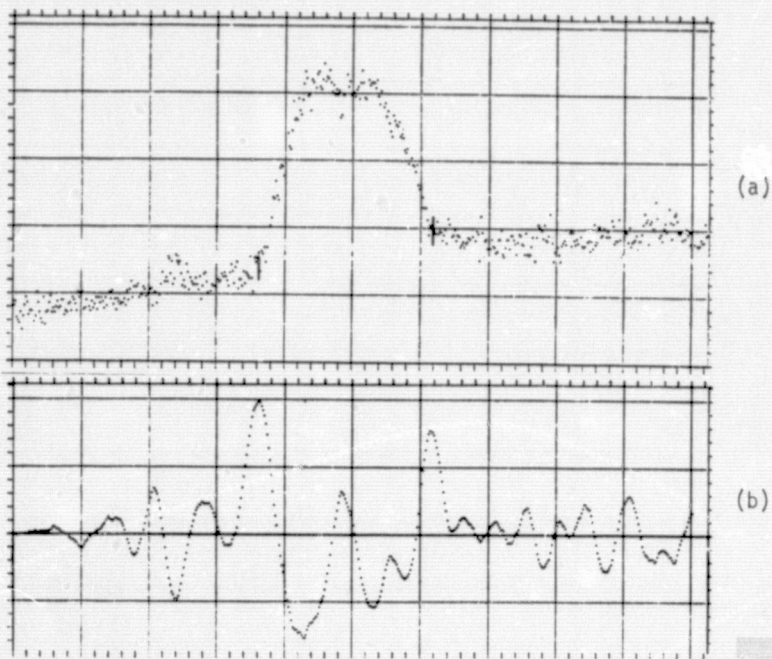
Figure 8. (a) Intensity profile with knee method edges marked.

(b) Second derivative plot of (a).
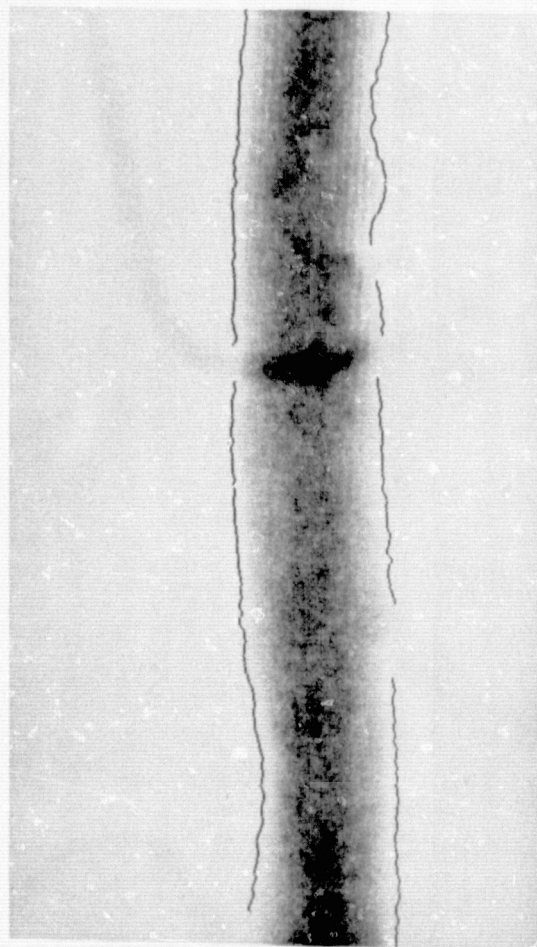
Figure 9. Femoral segment showing knee method edge detection.
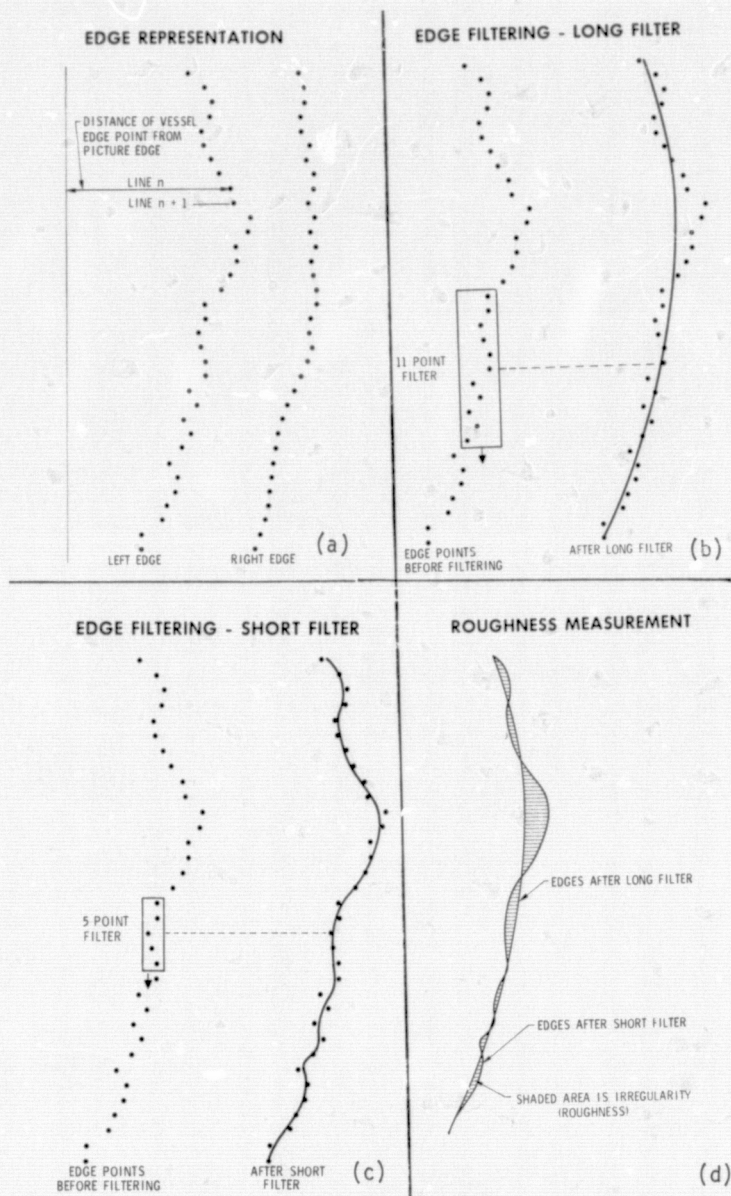
Figure 10. (a) Representation of vessel edges.

(b) Representation of a long filter applied to the left edges of Figure 10(a).

(c) Representation of a short filter applied to the left edges of Figure 10(a).

(d) Root-mean-square difference between short and long filtered edges define Roughness irregularity measure.
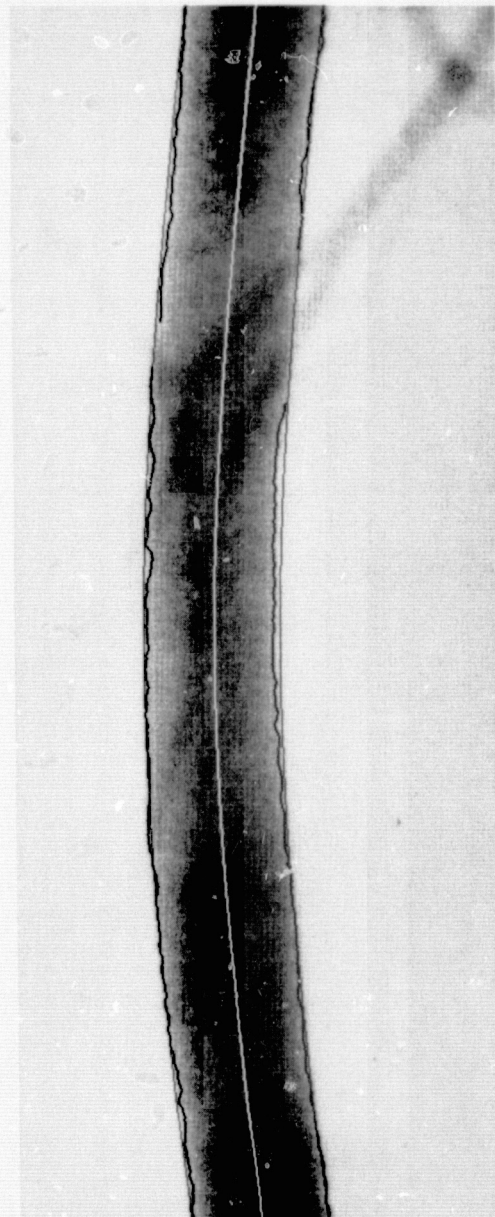
Figure 11. Example showing vessel midline and computer lumen lines derived for Lumen 90 measurement.

```
L90   331                                    242   L90
TL90  331                                    242   TL90
R81    16                                     18   R81
R97   135                                    112   R97
EL     44                                     45   EL

VG    175                                    157   VG
CC     42                                     35   CC


L90   705                                    309   L90
TL90  705                                    309   TL90
R81    25                                     20   R81
R97   125                                    158   R97
EL     43                                     50   EL

VG    243                                    179   VG
CC     39                                     49   CC


L90   446                                    593   L90
TL90  446                                    593   TL90
R81    17                                     15   R81
R97    83                                     66   R97
EL     47                                     41   EL

VG    199                                    211   VG
CC     27                                     22   CC


L90   193                                    586   L90
TL90  193                                    586   TL90
R81    15                                     20   R81
R97    90                                    105   R97
EL     46                                     43   EL

VG    130                                    226   VG
CC     29                                     33   CC


L90   210                                    390   L90
TL90  210                                    390   TL90
R81    24                                     13   R81
R97    34                                     40   R97
EL     47                                     53   EL

VG    165                                    171   VG
CC     13                                     14   CC


TOTALS LEFT   RT   AVE          LEFT   RT   AVE       WIDTH
 L90    421   454   438    EL    45    47    46    AVE 5423
 TL90   421   454   438    VG   185   191   188    MIN 4758
 R81     20    17    19    CC    32    31    32    MAX 6125
 R97    101   100   101                            TAP -.04
 B=112  V=216  CON=  727
```

Figure 12.   Standard femoral processing output image showing
             localized irregularity measures plus summary measures
             for entire vessel. The figure labels do not all
             correspond to measures described in text because of
             abbreviations used and because two labels were changed.
             Correspondence between the two is as follows:

Figure labels	Text label equivalent
R81	R(81,97)
R97	R(97,321)
EL	R(3,21)
VG	CEA
CC	CEC

B and V are the average background and mid-vessel
intensity values. CON is a measure of image con-
trast and is defined as

$$CON = \frac{V-B}{255-B}$$
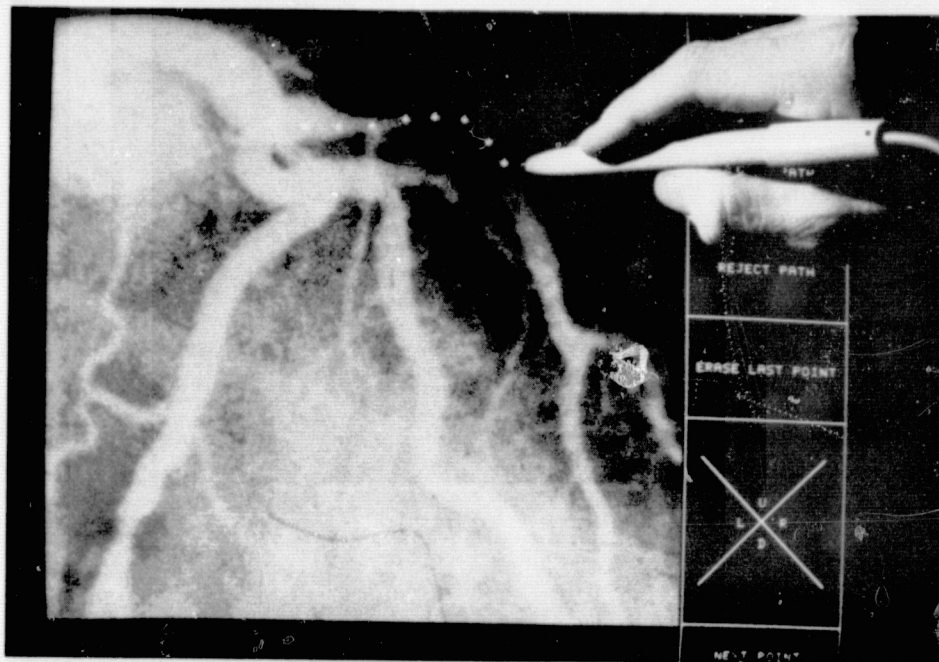
Figure 13.   Monitor display of coronary angiogram image showing use of sonic digitizer to provide approximate mid-vessel tracking path for computer edge finding program.
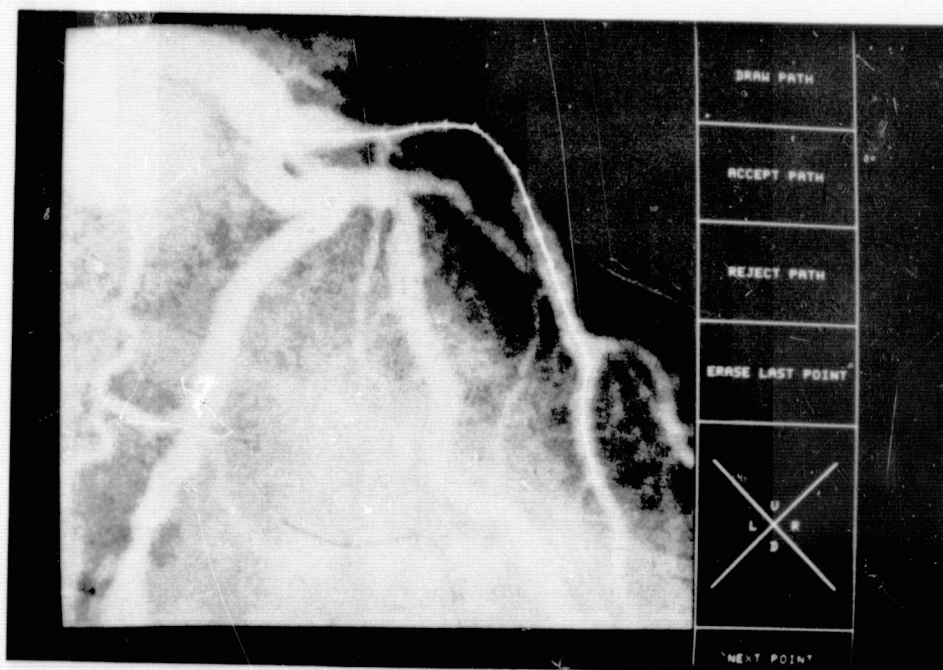
Figure 14.   Results of tracking path fill-in program.
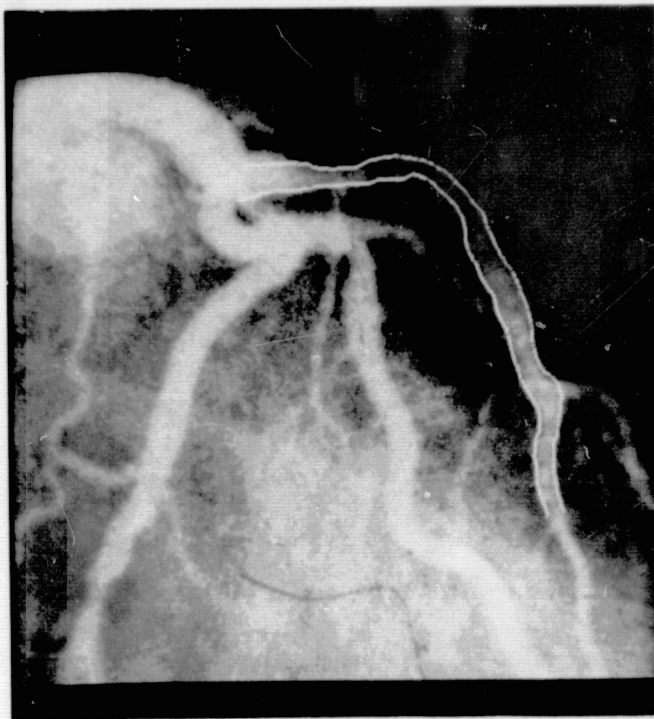
Figure 15.   Results of slope method edge tracking. Edges
were found along lines orthogonal to indicate
mid-vessel path.
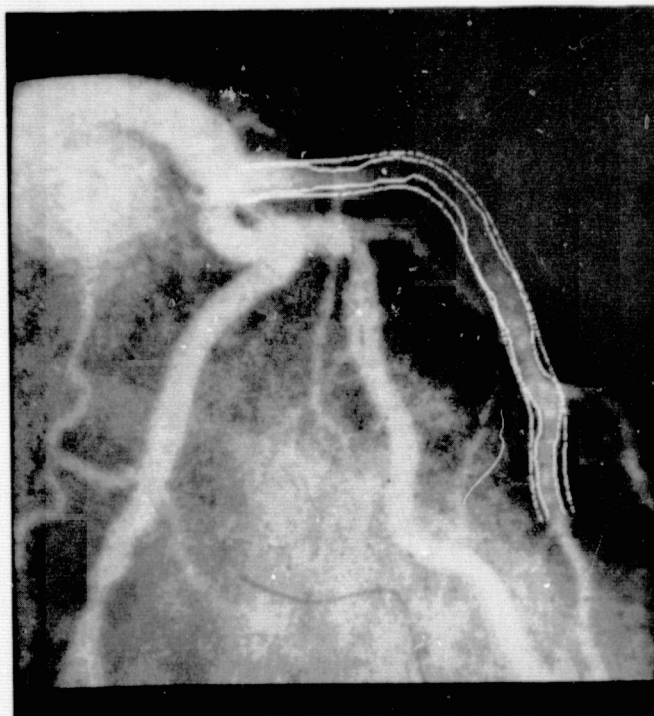
Figure 16.   Display of computer lumen reference edges
applied to coronary image.

SCAN REPLICATION - SLOPE EDGE METHOD

REPEAT PROCESSING OF SAME FILM

SCAN DATE	L90	TL90	R (81,97)	R (97,321)	R (3,21)	CEA	CEC
12/5/75	162	145	12	47	39	105	16
12/10/75	164	152	13	43	38	112	15
1/2/76	175	179	13	45	38	121	16
1/9	166	146	15	46	38	117	16
1/16	171	153	16	47	40	119	16
1/23	169	152	12	48	39	107	17
2/4	165	147	12	46	39	105	16
2/8	169	151	15	43	37	118	15
2/13	168	149	13	51	38	112	17
2/20	167	149	13	46	39	110	16
2/26	128	151	12	43	36	108	15
S.D.	11.9	8.8	1.4	2.3	1.1	5.5	0.7
MEAN	164.0	152.2	13.3	45.9	38.3	112.2	15.9
COEF VARIATION = S.D.×100 / MEAN	7.2	5.8	10.2	5.0	2.7	4.9	4.2

Table 1. Result of repeat scans of one film segment using slope edge method.

SCAN REPLICATION - KNEE EDGE METHOD

REPEAT PROCESSING OF SAME FILM

SCAN DATE	L90	TL90	R (81,97)	R (97,321)	R (3,21)	CEA	CEC
12/5/75	186	169	13	45	45	113	16
12/10	185	162	14	49	47	112	17
1/2/76	163	163	15	48	47	112	17
1/9	166	158	15	46	46	115	16
1/16	176	169	13	46	47	111	16
1/23	195	174	13	51	47	115	17
2/4	168	157	15	50	45	117	17
2/8	188	168	12	46	48	111	16
2/13	186	157	14	46	46	114	16
2/20	190	171	13	44	46	112	15
2/26	194	177	14	50	48	118	17
S.D.	10.9	6.6	1.0	2.2	1.0	2.3	0.6
MEAN	181.5	170.0	13.7	47.4	46.5	113.6	16.4
COEF VARIATION	6.0	4.0	7.0	4.7	2.1	2.0	3.9

Table 2. Result of repeat scans of one film segment using knee edge method.

EXPOSURE VARIABILITY - INDIVIDUAL MEASURES

MODERATE DISEASE PHANTOM	SMOOTH TUBE PHANTOM

MODERATE DISEASE PHANTOM — L90

% R \ KVP	75	80	85
10	534	553	539
20	506	520	493
30	491	493	505
40	474	489	473
50	486	480	508

S.D. = 23.4
COEF VAR = 4.6%

SMOOTH TUBE PHANTOM — L90

% R \ KVP	75	80	85
10	157	162	174
20	181	163	182
30	136	151	135
40	125	137	116
50	113	101	99

S.D. = 26.8
COEF VAR = 18.8%

MODERATE DISEASE PHANTOM — R (97,321)

% R \ KVP	75	80	85
10	147	159	142
20	147	141	148
30	147	144	148
40	141	146	144
50	143	143	144

S.D. = 4.3
COEF VAR = 2.9%

SMOOTH TUBE PHANTOM — R (97,321)

% R \ KVP	75	80	85
10	29	27	31
20	36	30	30
30	26	23	18
40	26	24	23
50	23	23	23

S.D. = 4.9
COEF VAR = 18.7%

% R = % RENOGRAFIN

Table 3. Exposure variability study results for L90 and R(97,321) for moderate disease phantom and smooth tube phantom. Three x-ray KVP values and five Renografin concentrations were used.

EXPOSURE VARIATIONS - MULTIPLE MEASURES

MODERATE DISEASE PHANTOM	SMOOTH TUBE PHANTOM

MODERATE DISEASE PHANTOM

CEA [1]

% R \ KVP	75	80	85
10	214	203	197
20	209	209	207
30	205	210	219
40	209	214	217
50	210	217	219

S.D. = 4.3
COEF VAR = 2.9%

SMOOTH TUBE PHANTOM

CEA

% R \ KVP	75	80	85
10	114	115	109
20	113	115	122
30	100	92	84
40	92	90	94
50	84	82	87

S.D. = 13.3
COEF VAR = 12.4%

CEC [2]

% R \ KVP	75	80	85
10	45	49	44
20	45	43	45
30	45	45	45
40	43	45	44
50	44	44	44

S.D. = 1.3
COEF VAR = 3.0%

CEC

% R \ KVP	75	80	85
10	11	11	12
20	13	12	11
30	10	9	8
40	10	9	9
50	9	10	9

S.D. = 1.4
COEF VAR = 13.5%

(1) COMPUTER ESTIMATE OF ATHEROSCLEROSIS
(2) COMPUTER ESTIMATE OF CHOLESTEROL

Table 4. Results of same experiment described in Table 3 for multiple measures, CEA and CEC.

SEQUENTIAL FILM PROCESSING

FOR EACH OF THE SIX PATIENTS, TWO SEQUENTIAL FILMS FROM
THE SAME EXAMINATION WERE PROCESSED

	COEFFICENT OF VARIATION	
MEASURE	SLOPE METHOD EDGES %	KNEE METHOD EDGE %
L90	8.3	5.7
TL90	8.8	3.9
R (81,97)	8.0	4.4
R (97,321)	8.8	6.1
R (3,21)	6.6	4.5
CEA	5.7	2.2
CEC	7.3	4.6

Table 5. Results of sequential film processing study. Film
pairs from same patient at same examination were
processed.

TEXTURE DIRECTED IMAGE COMPOSITING FOR FOCUS*

Ray J. Wall
Milan H. Karspeck

Jet Propulsion Laboratory
Pasadena, California

ABSTRACT

A considerable amount of physiological re-
search work involves light microscope examination
of histological specimens to determine microana-
tomical structure. However, high power light
microscopes are limited by their depth of view.
This paper describes a technique for reducing a
set of digitized optical sections to a single com-
posite image such that out of focus effects seen
in individual sections are reduced. The composite
image is derived from the optical sections on the
basis of a texture measure (gray level gradient).

I. INTRODUCTION

A considerable amount of physiological re-
search work involves light microscope examination
of histological specimens to determine microana-
tomical structure. However, light microscopes
have a characteristic which limit their usefulness.
A high power microscope objective lens has a field
of view several hundred microns in diameter but a
depth of view only a few microns thick. Further-
more, the focal field is not sharply bounded,
causing objects to disappear gradually as they
move out of the focal plane. Thus if an object is
non-planar the depth of view may be inadequate
such that structural detail will appear sharp in
some portions of the image while blurred in other
portions. This paper describes a digital image
processing technique that reduces the out-of-focus
portions of images of non-planar specimens. Our
approach uses optical sectioning to digitize the
specimen image in three dimensions. We segment
the image plane into subregions and decide for each
of the subregions, which of the optical sections
presents the sharper image. This decision is based
on a texture measure, image gray level gradient.
The mosaic of subregions defines a composite image.
The composite image presents sharper detail over
a larger extent than any of the component optical
sections. This procedure offers a reduction in
computational complexity over first constructing
a three-dimensional image and projecting onto a
plane to produce a two-dimensional picture for
viewing (Ref. 1).

II. OPTICAL SECTIONING AND IMAGE DIGITIZING

Specimen images are digitized by the Auto-
mated Light Microscope System (ALMS) at the Jet
Propulsion Laboratory (JPL). The ALMS is a com-
puter-controlled light microscope with 3-axis,
computer-driven stage motion and a dual-moving-
mirror type image-plane scanning digitizer. Stage
motion is controlled by stepping motors yielding
minimum stage sizes of 5 microns in the x and y
directions and 0.05 micron in the z direction

(axial). The image plane scanner (IPS) can digi-
tize either transmittance or optical density to
256 levels in a 1024 by 1024 point matrix. The
frame size in the image plane is 1 cm by 1 cm with
a 25 micron sampling aperture. The image plane
scales down to the specimen plane by the power of
the objective lens in use. We analyze slides at
40X (oil) magnification. At this power the 1024-
element sampling frame covers an area approximately
250 microns square in the specimen plane with a
.10 micron sample spot spacing. The microscope
and scanner are controlled by an IBM 1130 computer.
Before scanning, the logarithmic video amplifier
is calibrated such that gray level values are pro-
portional to the optical density of the specimen.
Images are then scanned line by line and written
on digital tape. This tape is analyzed on a PDP
11/45 in the Biomedical Image Analysis Facility
to produce the composite image. The image is
scanned repeatedly with z-axis stage position in-
cremented by 3 microns between scans. The auto-
mated optical sectioning and image digitizing pro-
cedure generates a digital picture on magnetic
tape.

III. PROCEDURE

Segmenting a picture into appropriate sub-
regions is generally an important initial step
toward further analysis. These regions are gen-
erally areas homogeneous with respect to a given
property or predicate. The case considered here
involves differentiating soft or blurred regions
from regions sharp in detail. In the analysis of
texture this is viewed as differentiating a soft
texture from a sharp texture. A standard texture
measure (Refs. 2 and 3) that reflects this quality
is based on taking differences between the gray
levels of pairs of nearby points. Gray level
differences in orthogonal directions are magnitude
summed to produce a measure that is reflective of
the gray level gradient. Each optical section was
divided into square subregions, 10 samples on a
side. The image gradient was computed for all
sample points that do not border the region and
the average gradient value computed. This is done
for each optical section and for the entire image.
A composite image is constructed from square sub-
regions by selecting the square subregion from
that optical section that produces the maximum
average gradient.

IV. EXPERIMENTAL RESULTS

Figures 1-3 are optical sections that depict
the morphology of the collagen network in the
interalveolar wall of the post-mortem human lung.
These three sections were taken at focal depth
separations of 3 microns. There are regions that

are sharp and distinct in one section and not in the others while there are regions that are not sharp in any of the three sections. This problem would have been remedied by taking more sections. Figure 4 shows the result of compositing the optical sections shown in Figures 1-3. Checkerboard effects can be observed in those regions that were not in focus in any of the regions. This is due to the lack of discrimination afforded by the texture of one plane versus another in these regions. Also we observe an occasionally misclassified cell in the areas where detail is the sharpest. This is due to a lack of edges within the 10 x 10 pixel region.

Figure 5 is the reconstruction map that indicates which of Figures 1-3 were used to generate the composite image. The non-planar variations in the tissue structure are more gradual and uniform than this map indicates. In order to reflect this information, the map was smoothed by a 5 x 5 uniform weight averaging filter. The result of filtering is shown by the reconstruction map of Figure 6. The map is now divided into larger more uniform regions as one would expect. Furthermore, the map reflects the overall convex structure of the alveolar wall. Figure 7 shows the composite image that was reconstructed according to the map of Figure 6. The checkerboard effects have been eliminated at the expense of the loss of sharpness over a small region of the image. This loss can probably be reduced by a more appropriate choice of filtering on the map of Figure 5.

V. FUTURE WORK

This work is preliminary in nature and is part of a continuing program of research into the morphology of the human lung. Future work in refining the image compositing procedure described here will include the following. First, the effectiveness of the gradient as a texture discriminant can be enhanced by ignoring image samples with comparatively low gradient values. Regions of relatively constant gray level should not be classified on the basis of gradient. Another texture measure (gray level variance) could detect these regions. If the variance is relatively low, then the reconstruction map for that region should be inferred from the neighboring regions.

REFERENCES

1. Weinstein, M. and Castleman, K.R., "Reconstructing 3-D Specimens from 2-D Images," Proceedings of the SPIE, Vol. 26, May 1971.

2. Rosenfeld, A., "Visual Texture Analysis: An Overview," TR-406, University of Maryland, August 1975.

3. Zucker, S.W., Rosenfeld, A. and Davis, L., "Picture Segmentation by Texture Discrimination," TR-356, University of Maryland, February 1975.
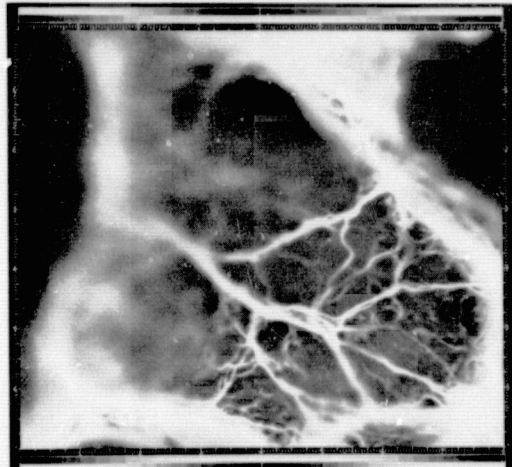
Figure 1. Optical section 1. This figure shows the collagen network embedded in the alveolar wall of a human lung.

Figure 2. Optical section 2. This section
was taken 3 microns from the
section shown in Figure 1.
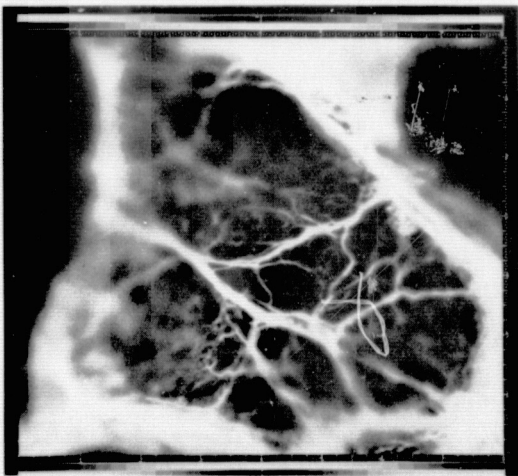
Figure 3. Optical section 3. This section
was taken 3 microns from the
section shown in Figure 2.
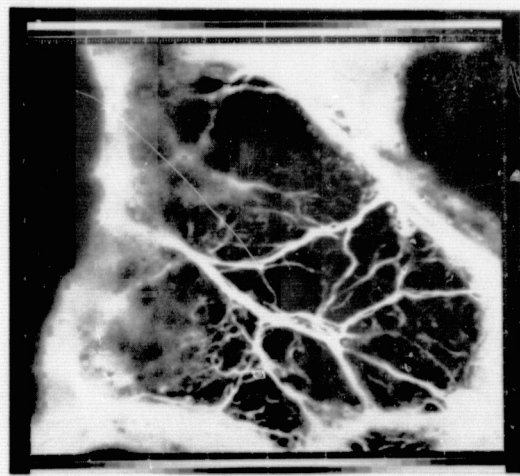
Figure 4. Composite image generated from
Figures 1, 2 and 3.

7-3
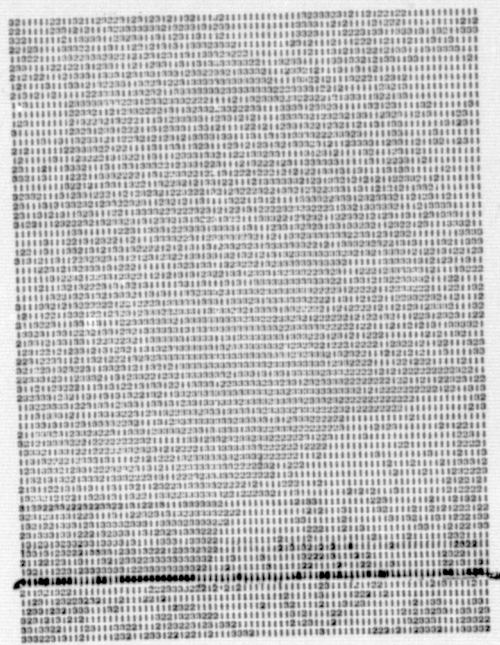
Figure 5. Image reconstruction map. Each number represents a 10 x 10 pixel region and the optical section. This map was used to generate Figure 4.
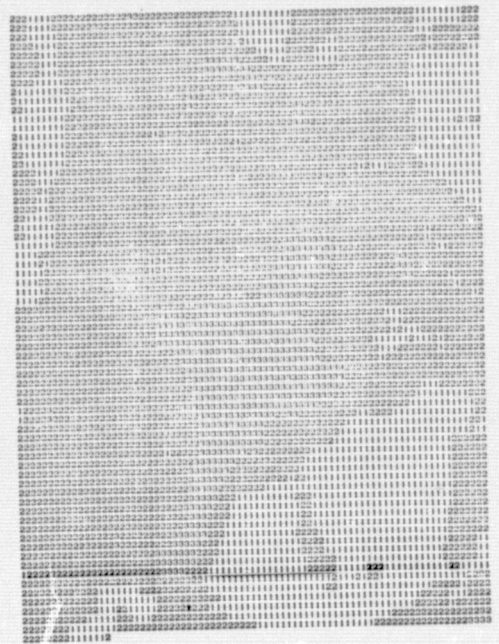
Figure 6. Filtered image reconstruction map. This map was derived from Figure 5 by the application of a 5 x 5 uniform weight smoothing filter.
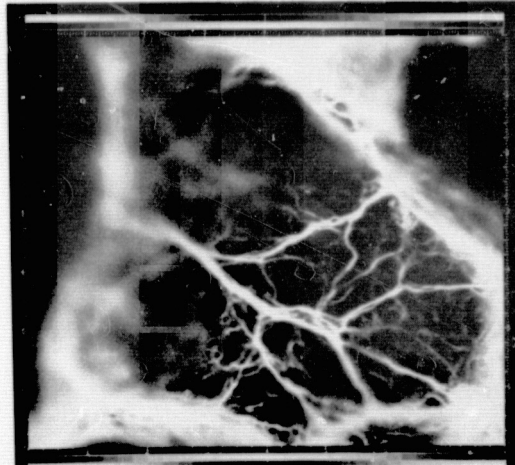
Figure 7. Composite image. This composite image was derived from the filtered reconstruction map shown in Figure 6.

REMOTE SENSING GEOLOGY: LANDSAT AND BEYOND

Alexander F. H. Goetz

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

ABSTRACT

Landsat data, particularly when computer image processing is applied, is a valuable tool in geologic mapping. In situ spectral reflectance studies yield data that can be analyzed to determine the optimum spectral bands, for geologic mapping purposes, to be used in future multispectral scanner systems. Future remote sensing efforts should be aimed at synthesizing data taken over a wide spectral range, from the UV to the microwave, for better discrimination and identification of materials not possible by the human observer.

I. INTRODUCTION

Remote sensing has been used in geologic mapping since the advent of the aerial photograph. After World War II, a sizable community of photogeologists emerged, using as their data source stereo photography from which surface maps and subsurface inferences could be made. This discipline plays a key role in petroleum exploration today.

Aerial photography is most useful because it displays the data in a context readily assimilated by the observer, i.e., in image form. However, the information recorded on film normally does not contain any data that could not readily be seen by a human observer viewing the scene from the same vantage point. This results from the fact that film and eye response are similar and are sensitive only to a very small portion of the electromagnetic spectrum. Remote sensing becomes a more valuable tool when data can be obtained in image form from parts of the spectrum outside the sensitivity range of the human eye and from a better vantage point. Remote sensing as discussed in this paper will be limited to imaging.

II. LANDSAT

The launch of ERTS-1 in 1972 signaled a new era in geologic remote sensing as well as remote sensing for other disciplines. The satellite offered a new vantage point, making visible relationships heretofore only inferred, extended imaging into the near infrared and, most importantly, made remote sensing data available to a much larger community of geologists than had ever used them before.

The most immediate use made of the Landsat images for geologic purposes was in the identification and mapping of large structures, particularly linear features which in some instances, with imagination, can be extended across continents (Carter, 1974). The origin of many of these features remains unclear, although some thorough field studies have been carried out to correlate lineaments in the images with features on the ground (Isachsen, 1973). Viljoen and Viljoen (1973) have demonstrated the value of Landsat for identifying structural domains in South Africa. Fischer and Lathram (1973) have identified a potential extension of the Umiat oil field in Alaska to the northwest, based on alignment of lakes and corroborative evidence in the similarity of magnetic and gravity data in both areas. These associations were unrecognized before the advent of the synoptic view from Landsat.

Mosaics of Landsat images have facilitated the understanding of plate tectonic relationships in areas where data were sparse or non-existent (Molnar and Tapponier, 1975). Halbouty (1976) provides an excellent review of the uses of Landsat data for petroleum and mineral exploration.

Although the MSS images have important value in structural interpretation and our understanding of this area is far from complete, the majority of research with Landsat has focused on the use of the multispectral imaging capability of the MSS. These data were new to most of the geologic community that had not taken part in previous aircraft remote sensing studies.

The basis for interest in the MSS spectral data stems from its potential for lithologic identification and detection of anomalies, be they structural, mineralogical or botanical, which can be used as indications of subsurface features.

Lithologic identification with the Landsat MSS is possible for only a few rock types (Rowan et al., 1974) because the wavelength region $0.5 - 1.1$ µm contains little unique spectral reflectance information. The reasons for this non-uniqueness will be discussed later.

Several methods of data manipulation are employed to extract the maximum amount of information available in the multispectral images. The methods all involve the use of digital images, since no radiometric degradation has taken place since acquisition, and the data are readily read by a computer. A number of digital enhancement techniques are discussed elsewhere in this compendium. In addition, many routines applicable to Landsat data are given by Goetz et al. (1975).

Digital enhancement is particularly useful to the geologist because he makes use of spatial information as well as color to discriminate rock units. Because unit boundaries are often

gradational, discrimination based on color alone can be misleading. Because rock units cannot be identified uniquely in the spectral region covered by Landsat, enhancement techniques are best applied to separate units by increasing the contrast among them in the image. A technique which has proved useful involves making ratios of band pairs and constructing color images from sets of three ratios (Billingsley et al., 1970; Goetz et al., 1971; Rowan et al., 1974). The ratio technique using one or more ratios has proved successful in the search for areas of hydrothermal alteration (Rowan et al., 1974, 1977). Alteration associated with ura.ium deposits (Spirakis and Condit, 1975; Offield, 1976; Vincent, 1977) and discrimination of individual basalt flows (Vincent and Thompson, 1972).

The supervised classification techniques have met with success in identifying potential porphyry copper deposits in Pakistan (Schmidt et al., 1975). However, only limited· success with supervised classification techniques have been reported by others (Siegal and Abrams, 1976) mainly because so much must be known about the area before training samples can be chosen. It would appear then that for reconnaissance mapping, image enhancement techniques such as ratioing, filtering, etc. would provide superior results, and for detailed work in an area which is at least partially well known, supervised classification techniques are suitable.

III. SPECTRAL REFLECTANCE

In the following the spectral reflectance of rocks and minerals in the region of 0.4 - 2.5 μm will be discussed. This particular region is chosen because on earth significant reflected solar energy is recorded. Short of 0.4 μm the solar contribution drops off rapidly and short of 0.3 μm atmospheric ozone shields the surface from UV radiation. Beyond 2.5 μm the major portion of energy coming from the earth's surface is in the form of emitted rather than reflected radiation.

In the visible and near IR region covered by Landsat, the reflectance spectrum of rocks and minerals is dominated by the transition metal cations, notably iron. Absorptions due to electronic transitions in iron are found in the 0.6 - 1.0 μm region. In the UV, strong charge transfer bands exist, the wings of which extend into the visible region, giving rise to an almost universally increasing reflectance with wavelength (Hunt and Salisbury, 1970).

At 1.6 μm no absorptions exist, and in this region the maximum reflectance spread in rocks is found, ranging from 10% to over 70%. Longward of 2 μm, overtones of fundamental lattice stretching and bending vibrations are found, giving rise to diagnostic OH and CO_3 ion absorption bands.

Fig. 1 shows several representative reflectance spectra of rocks taken in situ with the JPL Portable Field Reflectance Spectrometer (PFRS)

(Goetz et al., 1975). This instrument makes measurements of approximately 200 sq cm of the surface. A field standard (Fiberfrax) is measured at each site, and therefore measurements at different times are comparable. The discontinuities at 1.4 and 1.8 μm are the result of strong atmospheric water vapor absorptions, and very little energy reaches the earth's surface in these regions. Thousands of PFRS spectra have been acquired. They will be compiled and published at a later date.

As discussed previously, all the spectra exhibit an increasing reflectance with wavelength. In the Landsat spectral region the only materials which exhibit a diagnostic absorption band are those high in iron. Therefore limonitic rocks often found in areas of hydrothermal alteration can be uniquely identified in Landsat images after application of proper computer processing techniques (Rowan et al., 1974). The shape of the curves in this region is strongly dependent on the quantity and oxidation state of the iron (Salmon and Pillars, 1975).

In the region 2.0 - 2.4 μm diagnostic data for hydrous rocks and carbonates can be found. Clays exhibit an absorption at or near 2.2 μm while carbonates show a band at 2.35 μm as seen in Fig. 1.

An analysis was made of the spectra of altered and unaltered rocks, some of which are shown in Fig. 2 (Goetz et al., 1975). The object was to determine which spectral bands were most useful for identifying rocks associated with hydrothermal alteration. The continuous spectra were broken up into 13, 0.05 μm wide bands between 0.4 and 1.0 μm and 10 bands of 0.1 μm width in the 1.0 - 2.4 μm range. These bands were considered to be as narrow as practical for future scanner systems. Linear discriminant analysis was used to identify the bands in order of value for best separating altered and unaltered rocks. Discriminant functions or linear combinations of the bands were identified and plotted as shown in Fig. 3. Each axis is plotted according to a different discriminant function as applicable. The nine best bands (mostly beyond 1.0 μm) obviously separate the altered from the unaltered rocks well. But even three bands, 1.3, 1.6 and 2.2 μm, result in a far better separation than is achievable using the Landsat bands or any bands short of 1.0 μm.

IV. BEYOND LANDSAT

Fig. 4 is a graphic demonstration of the value of remote sensing in the wavelength region beyond 1.0 μm. The discriminant analysis technique can be applied to specialized problems in exploration or geologic mapping using existing field reflectance data as a base.

Not only are the absolute reflectances useful, but the ratios of spectral bands can be analyzed in a similar manner. Unfortunately, the number of combinations possible is 21!/2!, a very

large number. Some useful ratios of bands can be identified by inspection of the spectra. This method wa used in producing Fig. 4b. Fig. 4a is a color ratio composite of a Landsat image of a small area encompassing the Goldfield alteration zone, while Fig. 4b is constructed from ratios of spectral bands created by the JSC 24 channel MSS. Here the ratio of broad bands centered around 1.6 and 2.2 µm creates a large degree of separability among units, not as yet substantiated by field work, that are grouped together in the Landsat enhancement, Fig. 4a. An area directly to the south has been field checked, and the results from the image agree with the in situ spectral measurements (Abrams et al., 1977). This initial result points out that the correlation between the field spectral results and the multispectral images is positive and that analyses of this type of data will prove fruitful in developing specialized data gathering and analysis systems for geologic exploration.

Other wavelength regions beyond 2.5 µm, active and passive microwave imaging, and the correlation of these data with other geophysical parameters such as gravity and magnetics, were not covered in this paper. However, they too will play an important part in the synthesis of remote sensing data which will be required for future advances in the art of geologic mapping.

REFERENCES

1. Carter, W. D., "Tectolinear Overlay of the United States," U.S. Geol. Survey Open File Map, 1:5,000,000, 1974.

2. Isachsen, Y. W., Fakundiny, R. H., Forster, S. W., "Evaluation of ERTS-1 Imagery for Geological Sensing over the Diverse Geological Terranes of New York State," Symposium on Significant Results Obtained from the Earth Resources Technology Satellite-1, NASA SP-327, pp. 223-230, 1973.

3. Viljoen, M. J., and Viljoen, R. P., "ERTS-1 Imagery as an Aid to the Definition of the Geotectonic Domains of the Southern African Crystalline Shield," Symposium on Significant Results Obtained from the Earth Resources Technology Satellite-1, Vol. I: Technical Presentations: NASA SP-327, pp. 483-491, 1973.

4. Fischer, W. A., and Lathram, E. H., "Concealed Structures in Arctic Alaska Identified on ERTS-1 Imagery," Oil and Gas Jour., Vol. 71, pp. 97-102, 1973.

5. Molnar, P., and Tapponnier, P., "Cenozoic Tectonics of Asia: Effects of a Continental Collision," Science, Vol. 189, pp. 419-426, 1975.

6. Halbouty, M. T., "Application of Landsat Imagery to Petroleum and Mineral Exploration," Bull. Am. Ass'n. of Petroleum Geologists, Vol. 60, pp. 745-793, 1976.

7. Rowan, L. C., Wetlaufer, P. H., Goetz, A.F.H., Billingsley, F. C., and Stewart, J. H., "Discrimination of Rock Types and Detection of Hydrothermally Altered Areas in South-Central Nevada by Use of Computer-enhanced ERTS Images," U.S. Geol. Survey Prof. Paper 883, 35 p., 1974.

8. Goetz, A.F.H., Billingsley, F. C., Gillespie, A. R., Abrams, M. J., and Squires, R. L., "Application of ERTS Images and Image Processing to Regional Geologic Problems and Geologic Mapping in Northern Arizona," Jet Propulsion Laboratory Tech. Rept. 32-1597, p. 188, 1975.

9. Billingsley, F. C., Goetz, A.F.H., and Lindsley, J. N., "Color Differentiation by Computer Image Processing," Photographic Science and Engineering, Vol. 14, pp. 28-34, 1970.

10. Goetz, A. F. H., Billingsley, F. C., Head, J. W., McCord, T. B., and Yost, E., "Apollo 12 Multispectral Photography Experiment," Proceedings of the Second Lunar Science Conf., Vol. 3, pp. 2301-2310, MIT Press, 1971.

11. Rowan, L. C., Goetz, A.F.H., and Ashley, R. P., "Discrimination of Hydrothermally Altered and Unaltered Rocks in Visible and Near-infrared Multispectral Images," Geophysics, in press, 1977.

12. Spirakis, C. S., and Condit, C. D., "Preliminary Report on the Use of Landsat-1 (ERTS-1) Reflectance Data in Locating Alteration Zones Associated with Uranium Mineralization near Cameron, Arizona," U.S. Geol. Survey Open File Rept. No. 75-416, 20 p., 1975.

13. Offield, T. W., "Remote Sensing in Uranium Exploration," Amer. Inst. Min. Eng. (AIME) Proceedings of Meeting, Vienna, Austria, March 30-April 2, 1976, in press, 1976.

14. Vincent, R. K., "Uranium Exploration with Computer-processed Landsat Data," Geophysics, in press, 1977.

15. Vincent, R. K., and Thomson, F. J., "Rock-type Discrimination from Ratioed Infrared Scanner Images of Pisgah Crater, California," Science, Vol. 175, pp. 986-988, 1972.

16. Schmidt, R. G., Clark, B. B., and Bernstein, R., "A Search for Sulfide-bearing Areas Using Landsat-1 Data and Digital Image Processing Techniques," in Proceedings of the NASA Earth Resources Survey Symposium, NASA, TM X-58168 (JSC-09930), Vol. I-B, pp. 1013-1038, 1975.

17. Siegal, B. S., and Abrams, M. J., "Evaluation of Classification Techniques Applied to Geologic Mapping Using Landsat Data," in Proc. of Am. Soc. of Photogrammetry, 42nd

Annual Mtg., Washington, D. C., pp. 543-544, 1976.

18. Hunt, G. R., and Salisbury, J. W., "Visible and Near-Infrared Spectra of Minerals and Rocks: I Silicate Minerals," Modern Geology, Vol. 1, pp. 283-300, 1970.

19. Salmon, B. C., and Pillars, W. W., "Multi-spectral Processing of ERTS-A (Landsat) Data for Uranium Exploration of the Wind River Basin, Wyoming: A Visible Region Ratio to Enhance Subtle Alteration Associated with Roll-type Uranium Deposits," Environ. Res. Inst. of Mich., Rept. 110400-2-F, Ann Arbor, Mich., 1975.

20. Goetz, A. F. H., Siegal, B. S., and Rowan, L. C., "Quantitative Spectral Techniques and Computer Image Processing as Applied to Lithologic Mapping," Proceedings of 14 IEEE Conference on Decision and Control, Dec. 10-12, Houston, Texas, pp. 412-413, 1975.

21. Abrams, M. J., Goetz, A. F. H., Kahle, A. B., Ashley, R., Rowan, L. C., "Lithologic Mapping by Spectral Imaging in the $0.4 - 2.5$ μm Region: Ralston Mining District, Nevada," submitted to Geology, 1977.

ACKNOWLEDGEMENT
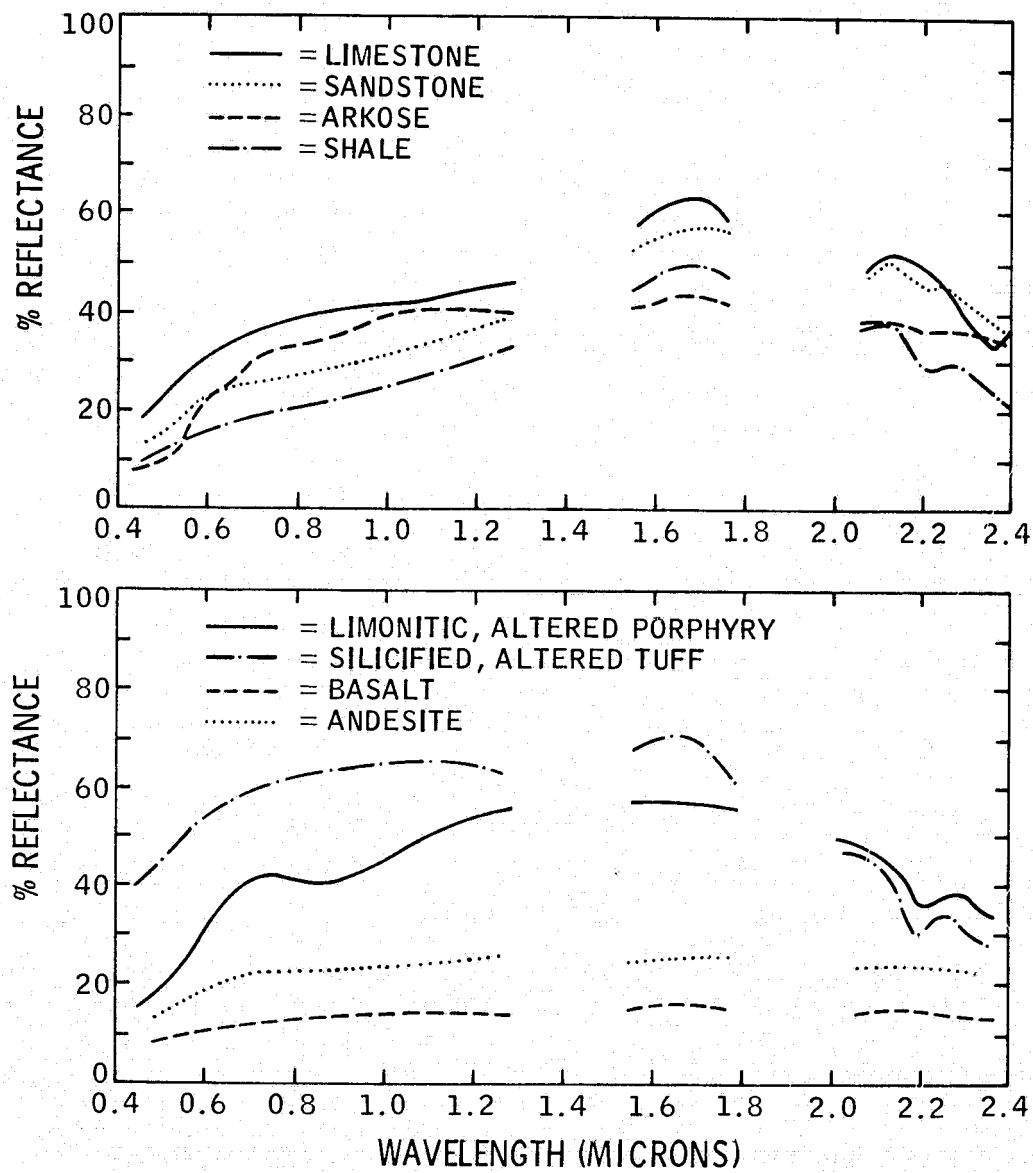
Fig. 1. Field reflectance spectra of a representative group of
igneous and sedimentary rocks.
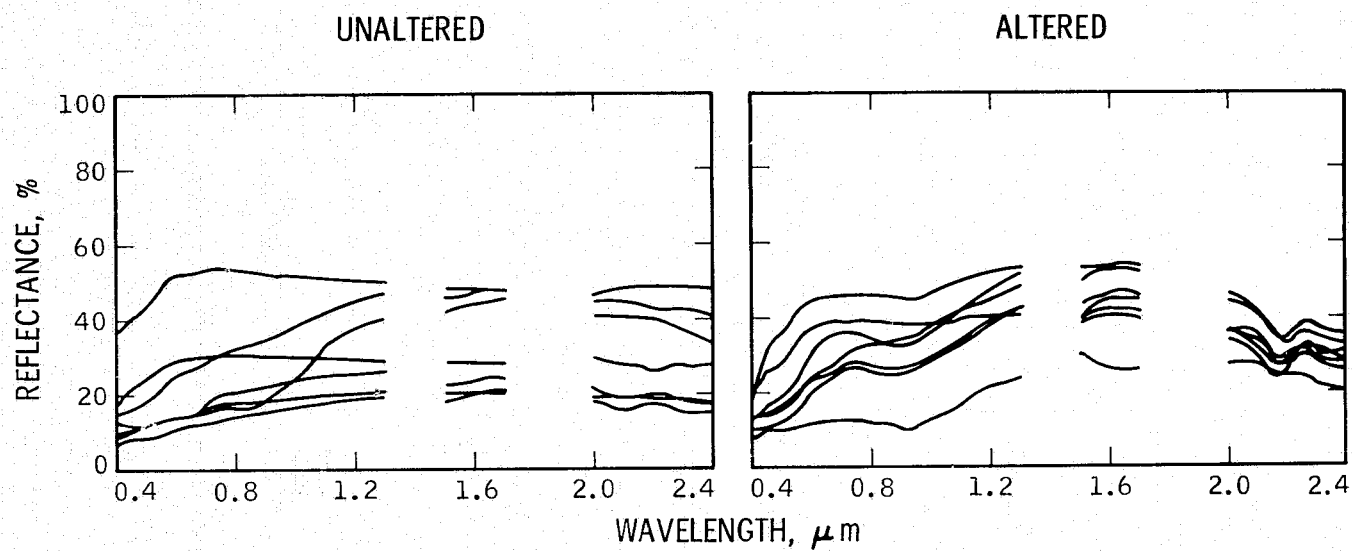
UNALTERED   ALTERED

Fig. 2.   Field reflectance spectra of altered and unaltered rocks.
Note the strong 2.2 μm band in all the spectra of altered
rocks.

ERTS BANDS

0.4 - 1.0 μm
13 BANDS

1.6, 2.2, 1.3 μm

9 BEST
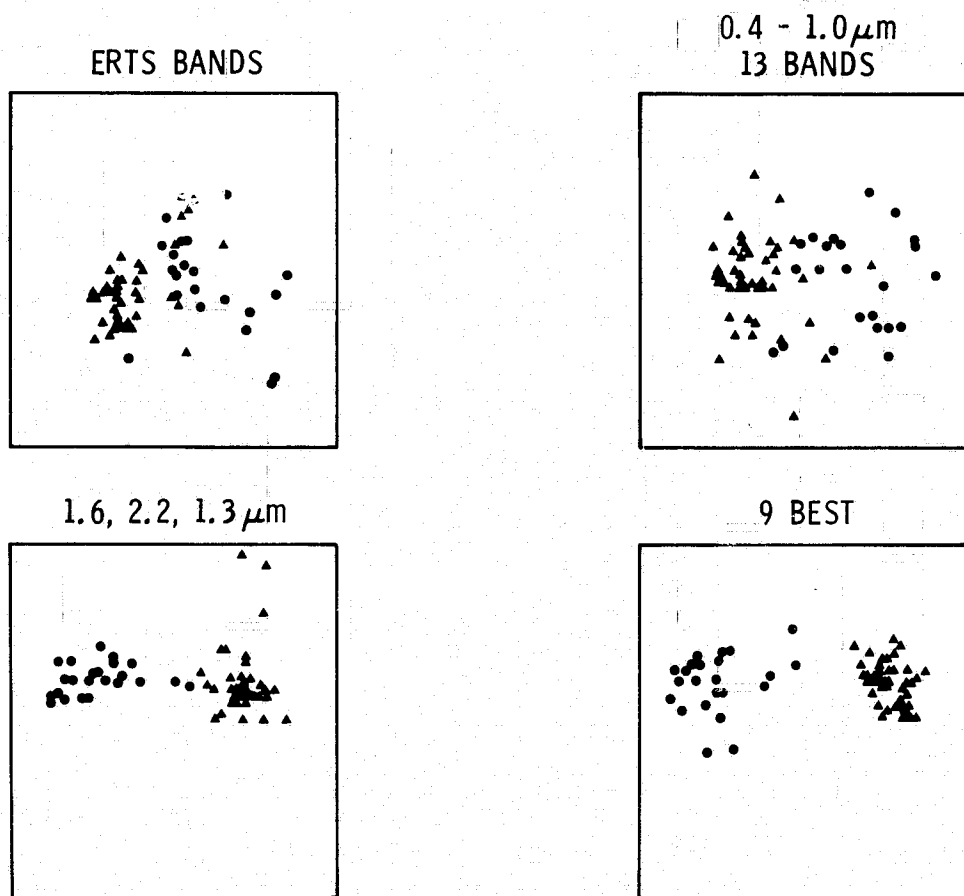
Fig. 3. Stepwise linear discriminant analysis showing
separability of a group of altered and unaltered
spectra, a sample of which is shown in Fig. 2.
Circles represent the altered materials; triangles,
unaltered materials. The axes represent linear
combinations of the reflectance values in the
bands in question.
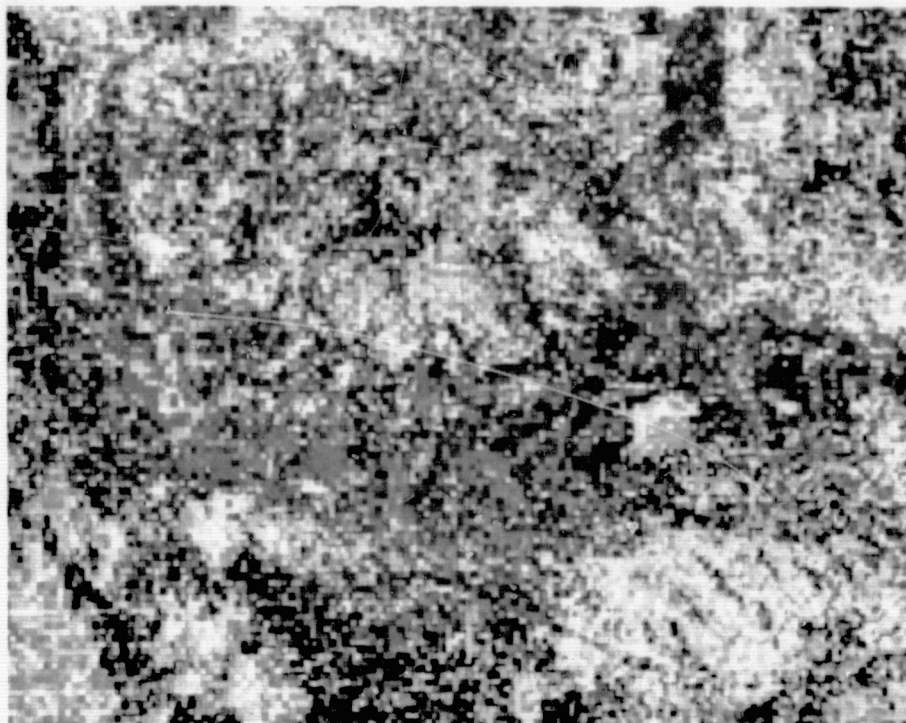
Fig. 4a. Landsat color ratio composite of a 15 x 20 km area surrounding the Goldfield, Nevada mining district. The colors were constructed by assigning blue to 4/5, green to 5/6, and red to 6/7.
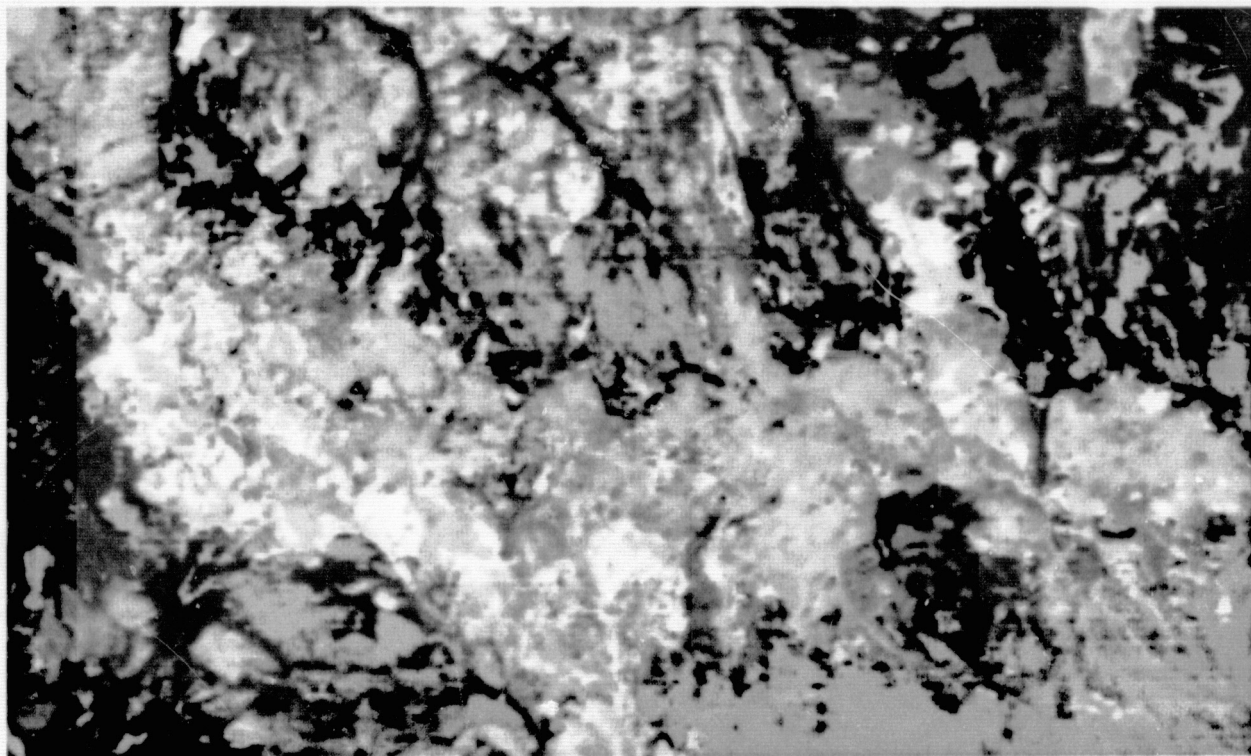
Fig. 4b. JSC 24 channel scanner color ratio composite of the same Goldfield area made from ratios of bands 0.6/1.0, 1.6/2.2, 1.6/0.48 displayed as blue, green, and red respectively. The wavelengths in micrometers represent band centers. The data has been smoothed to represent the 30 m IFOV anticipated for the Thematic Mapper on Landsat D.

INTEGRATION OF SOCIOECONOMIC DATA AND REMOTELY SENSED IMAGERY FOR LAND USE APPLICATIONS

Nevin A. Bryant, Ph. D.

Jet Propulsion Laboratory
Pasadena, California

ABSTRACT

Metropolitan regions have in recent years become increasingly impacted by competing demands for natural and economic resources from various political and commercial constituencies. In an effort to better plan for resource utilization, many state and local governments are establishing geographic information systems which can manage spatially referenced data, perform certain types of spatially-oriented processing, and are current and comprehensive. Geographic information systems, which have progressed roughly at the rate of advancement in computer technology, seek to capitalize on the synergism inherent in being able to automatically compare a variety of socioeconomic, environmental, and land use data sets for the same point on the ground.

Polygon overlay and grid cell information systems assess data for selected areas, but their data files are time-consuming to generate and frequently costly to process. Updating of land use changes for such systems may become prohibitively expensive. An alternative system, IBIS (Image Based Information System) is presented using actual cases for illustration. IBIS makes use of digital image processing techniques to interface geocoded data sets and information management systems with thematic maps and remotely sensed imagery. The basic premise is that geocoded data sets can be referenced to a raster scan that is equivalent to a fine mesh grid cell data set.

Two cases illustrating the use of IBIS to integrate socioeconomic data and remotely sensed imagery are presented. In the first case, thematically classified LANDSAT imagery covering Orange County, California, is aggregated by census tracts and municipalities for two dates. The Orange County administrative office, which had contracted for a polygon overlay system tabulation of land use derived from aerial photography for 1972, was interested in comparing the utility of LANDSAT-derived thematic mapping for the same time period and assessing the cost-effectiveness of a LANDSAT-derived land use update for 1975. In the second case, thematically classified LANDSAT imagery over the City of Los Angeles was aggregated by census tracts. Here, May and September 1975 LANDSAT frames were registered and combined digitally to permit incorporation of seasonal differences to enhance the thematic classification. Furthermore, thematic classification was performed separately on various regions of the city having similar age and land cover characteristics, thereby reducing the chance of misclassification attributable to signature extension. Regionalization of the city prior to thematic classification was accomplished by using selected portions of the polygon reference image of census tracts as a binary mask applied to the raw LANDSAT imagery.

INTRODUCTION

Metropolitan regions have in recent years become increasingly impacted by competing demands for natural and economic resources. As the value of land and the cost of construction materials has increased, monitoring urban change and modelling urban trends has moved from being of purely academic and long-range planning interest to becoming an important factor in economic decision-making for the near term. Scarcity of natural resources, particularly energy, water and land, have intensified the conflict between commercial and environmental interest groups. For the growing metropolises of the South and West concerns focus on the need to accurately project future energy demand and conversion of agricultural land, while the predominant concern of stable cities of the East and North is the more efficient utilization of existing resources and discovering ways to revitalize the inner city.

In an effort to better plan for resource utilization and more fairly assess the impact of alternative proposals for land use, many state and local governments are expanding their use of remote sensing and incorporating the interpreted results of remote sensing into geographic information systems. For instance, changes in land use noted over a two year period may be used to help update decennial Census statistics and predict future growth trends. Characteristically, no individual data source, such as a census or survey, can provide answers to the variety of problem addressed by decision makers. The cost of comprehensive surveys is usually too expensive for local governments, whereas the combination of a variety of existing data sets (e. g., building permits) and remote sensing products frequently provide estimates that fall within the statistical error inherent in survey techniques.[1] Geographic information systems which have progressed roughly at the rate of advancement in computer technology, seek to capitalize on the synergism inherent in being able to automatically compare a variety of socioeconomic, environmental, and land use data sets for the same point on the ground.

Geographic information systems should satisfy four basic criteria if they are to be useful: 1) They should provide specific point locations, as well as area locations of data; 2) They should provide for variable aggregation (sub-setting) of data; 3) They should provide a method for representing spatial arrangements; and 4) They should be able to interface with mathematical and statistical programs which can be called as needed to aid in the analysis of spatially-oriented data.[2] Practitioners of the art of geocoded systems design have progressed with varying degrees of success towards the goals outlined. As a rule, generalized systems have only rudimentary data manipulation capability (i. e., status updating and interrogation by

area), while highly specific and specialized systems have progressed further with modelling applications.[2]

In response to the desire to access data for selected areas, polygon and grid cell geocoded information systems have been developed. Such systems rely on the tabular formatting of the input data, a costly and time-consuming process. Often the system falls into disuse because the updating of major segments of the data base becomes prohibitively expensive. In response to the desire to provide up-to-date resource information, investigators have studied the feasibility of applying LANDSAT and high altitude photo imagery to natural resource mapping. The consensus which appears to be evolving is that, while remotely sensed imagery can provide timely coverage and sufficiently accurate maps using ADP techniques, the end product is still a map if it cannot interface directly with an existing geocoded information storage and retrieval system.[4]

INTERFACING GEOGRAPHIC INFORMATION SYSTEMS AND REMOTE SENSING DATA

Data Management Considerations

The ease with which an agency can establish a geographic information system is constrained by the level of detail and the computer technology available. With geocoded data, there is a dramatic increase in file size with every added variable and increased resolution. Thus, a parcel level system for the City of Los Angeles consists of over 800,000 records, while a census tract level system has less than 1,000. The use of newer generation computers has only moderately improved the overall operation, as the major efforts are involved in both file generation and editing and computer software architecture. Frequently, urban information systems have become underutilized because of high manpower requirements for the initial encoding of data and the heavy expense involved in updating. Land use characteristics have consistently been the most expensive data to incorporate in information systems, primarily because there has been no subsidy available to the user in recording and preparing resource inventory data, such as occurs with demographic, economic, health, and assessor data sets.[5]

There are a number of ways to encode spatial data.[6] It is possible, however, for illustration purposes, to dichotomize referencing systems into nominal and ordinal, and data types into tabular, graphical, and image. Ordinal systems reference data by the actual geographical coordinate values. Thus, natural resources such as forests, rivers, crop land, and geologic formations are mapped with selected identifiers (total area, boundary, centroid) referenced to latitude and longitude or other selected geographic coordinate system. Nominal systems are "name referencing," i.e., data or information is referenced to a name-designating system. Any district-based referencing convention, such as census tract, sewer district, township, or transportation zone, assumes the operator knows where each administrative area is located and leaves the analysis of contiguity-effects, etc., up to the individual.[7] Problems invariably arise during the computational processing involved in conversions

between ordinally and nominally referenced data and have, with the exception of the Census DIME (Dual Independent Map Encoding) File system, forced the conversion of all data to one reference system or the other.[8]

Of equal concern to the designer of a geographic information system is an incorporation of the various geocoded data types to assure adequate "data capture". It is only through the integration of tabular, graphical, and image data types that geographic information systems achieve the synergistic impact required to offset their initial cost. For instance, tabular files may keep records of individual weather stations, while graphical files record elevation contours, and image data sets the distribution of land use. The combination of all three data types would provide analysts with the variety of spatial data needed to model alternative sources of non-point source air pollution. The need to deliver a uniformly encoded result to the user has been a key element to changes that have occurred in both geocoding approaches and computer system architecture applied to geographic information systems.[9]

Approaches to Geocoding

There exist three principal geocoding architectures, which have evolved from simple grid cell systems to the complex polygon methodologies, and most recently include the image raster data type (see Fig. 1). From several recent reviews, it is evident that the goals of geographic information systems have become more ambitious as faster computer systems evolved and peripherals become more sophisticated.[10] Grid cell methods served the need to retrieve geo-located data and generate maps through the cross tabulation of variables encoded within a particular cell. Several important drawbacks reduce the overall flexibility of grid cell systems: (a) Their spatial resolution is only as accurate as the grid cell size (usually ranging from one acre to ten square miles); (b) The systems permit the referencing of data in either a nominal or ordinal manner, never both; (c) The need for manual encoding of the input data files has made updating difficult and even prohibitively expensive and effectively limited the spatial resolution of grid cells to satisfy the need to achieve regional coverage.

In response to the failings of grid cell geocoding, polygon systems grew as electronic coordinate digitizers became generally available. Polygon geocoding formats effectively solved the spatial resolution dilemma inherent with grid cell
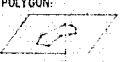
METHOD	COST CONSIDERATION	OVERALL FLEXIBILITY
GRID CELL:	MANUALLY OPERATED	SPATIAL RESOLUTION POOR, UPDATING DIFFICULT
POLYGON:	EXPENSIVE FOR LARGE DATA SETS	CERTAIN OPERATIONS PROHIBITED
IMAGE RASTER:	REQUIRES IMAGE PROCESSING TECHNOLOGY	NEITHER SCALE NOR DATA FORMAT DEPENDENT

Figure 1. Approaches to Geocoding

formats, while coordinate digitizing hardware has permitted rapid encoding of data. The most important achievement, however, was the integration of nominal and ordinal referencing in the Census DIME file methodology.[11] Despite these significant achievements, polygon geocoding systems have left the problem of ordinal data updating unresolved and created new challenges inherent in their graphical data structure. These problems include: (a) Considerable computational expense associated with file editing; (b) Complex topological architectures to achieve efficient data extraction from any given areas;[12] and (c) Large investments in computer systems to achieve polygon overlay of separate files for encoding ordinal data into nominal encoding formats (e.g., acreage of land use for each census tract). Many of these constraints can be mitigated by the use of an image raster encoding procedure and application of digital image processing algorithms to implement geographic information system analyses.

Digital image processing techniques can be applied to interface existing geocoded data sets and information management systems with thematic maps and remotely sensed imagery. The basic premise is that geocoded data sets can be referenced to a raster scan that is equivalent to an ultra-fine mesh grid cell data set, and that images taken of thematic maps or from remote sensing platforms can be converted to a raster scan. A major advantage of the raster format is that x, y coordinates are implicitly recognized by their position in the scan, and z values can be treated as Boolean layers in a three-dimensional data space. Such a system should permit the rapid incorporation of data sets, rapid comparison of data sets, and adaptation to variable scales by resampling the raster scans.

The development and system characteristics of an image based information system (IBIS) that makes use of digital image processing techniques has been described by Bryant and Zobrist in June 1976.[13]

REMOTELY SENSED DATA INPUT

A major impetus to the creation of geographic information systems has been the desire to integrate land resource inventory data, derived from remote sensing imagery, with other geocoded statistics. The costs of acquiring and encoding remotely sensed data has frequently spelled the demise of an agency's plans to create a geographic information system, while the lack of funds for land resource inventory updates to match other statistical updates has frequently left systems in disuse.[14] A recent survey of the status of land use mapping and inventories for the Southern California region illustrated both the problems of updating and lack of uniformity for classification and spatial area aggregation.[15]

For several years NASA has sought to partially alleviate the financial burden of procuring remote sensing imagery useful to resource inventory assessment through the provision of high altitude and LANDSAT imagery at minimal cost to the user. The ability of IBIS to incorporate LANDSAT imagery directly in a geocoded system framework should help resolve much of the present impasse.

Monitoring Potential

The ultimate goal for NASA support to geographic information systems, periodic land resource inventory updates for statistical areas and semiannual maps locating areas that have undergone change, has yet to be achieved.[16] The technological limits of existing automated data processing systems and the resolution capability of the current LANDSATs have permitted only partially accurate classification and statistical results. Particularly in the urban scene, abrupt changes in land use occur over short distances and often several cover types occur within a particular class of land use. In a short distance, one may encounter asphalt freeways, residential areas with light colored roofs adjacent to those with dark colored roofs, and then concrete freeways.

The results of experiments completed to date show that the current generation of LANDSATs have a spatial resolution that can provide USGS Level II land use statistics of sufficient accuracy when aggregated at the census tract level or higher under closely edited test conditions.[17] Higher spatial resolution simulations of LANDSAT-C follow-on missions are proving that more thematic classes can be differentiated; and, more significantly, that more accurate acreage estimates by statistical analysis areas (e.g., census tracts) are to be expected.

CASE STUDY DESCRIPTION

Two cases illustrating the use of IBIS to integrate socioeconomic data and remotely sensed imagery follow. In the first case, thematically classified LANDSAT imagery covering Orange County, California is aggregated by census tracts and municipalities for two dates. In the second case, thematically classified LANDSAT imagery over the City of Los Angeles was aggregated by census tracts.

CASE 1: Orange County, California

Orange County is divided into some 315 census tracts and 26 municipalities. For general purposes of planning and management, it was desired to tabulate land use by census tracts and municipalities for 1972 and 1975 using LANDSAT data. A mapping of land use from aerial photography and tabulation of statistics by census tracts using polygon techniques had been performed in 1972. The Orange County Administrative Office was interested in comparing the utility of LANDSAT-derived thematic mapping for the some time period and assessing the cost-effectiveness of a LANDSAT-derived land use update for 1975.

Three sets of input data were needed for this case: 1) district maps of census tracts and municipalities digitzed and stored as two files on a computer tape, 2) LANDSAT images of the same area, and 3) user-identified training sites which outlined sample areas known to contain land cover types of interest. Processing steps are shown in Figure 2 with the input data at the top and desired output at the bottom. The sequence of processing steps were:
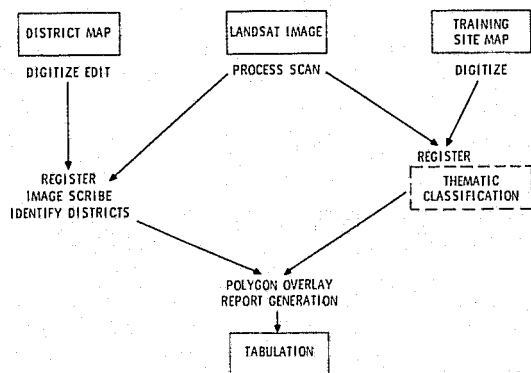
Figure 2. Processing Steps

1. Convert the polygon district files of census tracts and municipalities into images as shown in Figure 3 (automated).

2. Apply rubber-sheet stretching to achieve registration of census tract and municipality boundaries to the LANDSAT image as shown in Figure 4 (automated, manual selection of tie points).

3. Extract the region of interest from a LANDSAT digital image and convert to a format suitable for classification algorithms (automated).

4. Digitize the training site map on a coordinate digitizer (manual, magnetic tape output).

5. Repeat step 2 for the training sites.

6. Apply mathematical classification[18] to the LANDSAT image and training sites to produce a thematic classification of the entire region of interest as shown in Figure 5 (automated).

7. Identify the census tracts and municipalities in the administrative district images by painting each with a unique pixel value or grey scale as shown in Figure 6 (automated).

8. Perform polygon overlay, summing the number of pixels of each thematic class for each of the colors painted in an administrative image. By overlaying the two images, each location corresponds to a pair of pixels, one representing a district and the other a thematic class. Polygon overlay then becomes a trivial histogramming operation on these pixel pairs. (automated)

9. Convert the pixel histogram into a report with columns of land use by acres and percentages shown in Figure 7. This involves a number of transformations in the data management and analysis portion of IBIS. (automated)

10. Repeat steps 6 through 9 for the second LANDSAT image. (automated)

The theme class image is a useful product since it helped the Orange County personnel compare visually the distribution of land use in the region as it was interpreted from the satellite and aerial photography. But for purposes of analysis, the data needed to be aggregated according to census tracts and municipalities, as all other data collected within the county are aggregated in these districts. The IBIS system gave both products as well as other material that permitted Orange County personnel to check the system accuracy at each procedural step.

CASE 2: City of Los Angeles

A frequently used unit of aggregation in metropolitan areas is the census tract, defined by the U.S. Bureau of the Census. Census data are available by that unit and data gathered by other district conventions (e.g., fire district, police district) are often crosstabulated to census tract so that analysis can be performed. Moreover, annual statistical updates in metropolitan regions are made at the census tract level, but are not made for a smaller unit of aggregation. In this IBIS application, it was desired to tabulate land use by census tract. Los Angeles is a large case involving thousands of square miles and over 1500 tracts. The tracts have been coordinate digitized and edited by the Census Bureau and are available at low cost on magnetic tape.

Referring back to the processing steps for Case 1, the first step uses a tape handling routine to convert the census tape contents to IBIS polygon file format, then a polygon rotation routine converts the longitude-latitude coordinates to line-sample coordinates in approximate registration to the LANDSAT image. Rubber-sheet registration using about thirty tie points then obtains precise registration.

Except for the steps leading to thematic classification, the Los Angeles case is less costly than the Orange County case and demonstrates the economies of scale inherent in the IBIS system. In an effort to achieve greater discrimination among cover types, May and September 1975 LANDSAT frames were registered and combined digitally to permit incorporation of seasonal differences. Furthermore, thematic classification was performed separately on various regions of the City having similar age and land cover characteristics, thereby reducing the chances of misclassification attributable to erroneous signature extension. Regionalization of the City prior to thematic classification was accomplished by using selected portions of the administrative image of census tracts as a binary mask applied to the raw LANDSAT imagery.

Figure 3a. Orange County Census Tract File
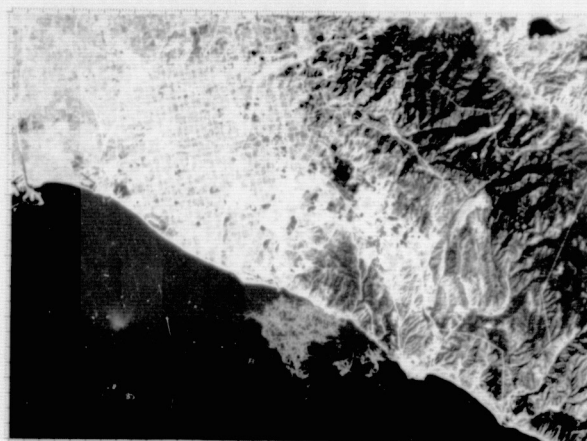
Figure 4a. Orange County Census Tracts
Registered to Image

Figure 3b. Orange County Municipal
Boundary File
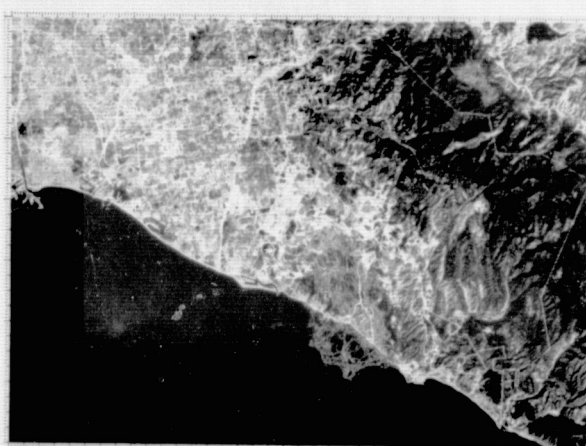
Figure 4b. Orange County Municipal Boundaries
Registered to Image

BIBLIOGRAPHY

1. "FY76 Summary Report on the Thematic
Mapper Simulation in Metropolitan Regions"
Typescript. Earth Resources Applications
Group, Jet Propulsion Laboratory, Pasadena,
California, 12 July 1976.

2. Kohn, C. F., "The 1960's: A Decade of
Progress in Geographical Research and
Instruction," Annals of the Association of
American Geographers, Vol. 60, No. 2
(June 1970), p. 215.

3. Dyer, H. L., et al., Information Systems for
Resource Management and Related Applica-
tions, Argonne National Laboratory, for the
Office of Land Use and Water Planning,
2 volumes, 1975.

4. Eastwood, L. F., et al., Preliminary Needs
Analysis Report: Program on Earth Observa-
tion Data Management Systems, Center for
Development Technology, Washington Uni-
versity, St. Louis, MO, December 31, 1975.

5. Miller, W. R., A Survey of Geographically
Based Information Systems in California,
Intergovernmental Board on Electronic Data
Processing, State of California, May 1975;
Power, M. A., Computerized Geographic
Information Systems: An Assessment of
Important Factors in Their Design, Opera-
tion, and Success, Center for Development
Technology, Washington University,
St. Louis, MO, December 31, 1975.

6. Deuker, K. J., "A Framework for Encoding
Spatial Data," Geographical Analysis, Vol. 4
(1972), pp. 98-105; Deuker, K. J., "Urban
Geocoding," Annals, Association of American
Geographers, Vol. 64 (1974), pp. 318-325;
Werner, P. A., "National Geocoding,"
Annals, Association of American Geograph-
ers, Vol. 64 (1974), pp. 310-317.

7. Cooke, D., Storage and Retrieval of Geo-
graphic Data, URISA, SIG/GBF Technical
Session, Atlantic City, NJ, 1973.

ORANGE COUNTY LAND USE
NOVEMBER 22, 1972

☐ LOW DENSITY RESIDENTIAL		☐ AGRICULTURAL	
☐ MED./HIGH DENSITY RESIDENTIAL		☐ GRASSLANDS	
☐ COMMERCIAL, INDUSTRIAL AND INSTITUTIONAL		☐ CHAPARRAL	
☐ FLOOD CHANNELS AND EXTRACTIVE		☐ WETLANDS	
☐ UNDEVELOPED URBAN		☐ WATER	
☐ GREEN SPACE		— CENSUS TRACTS	
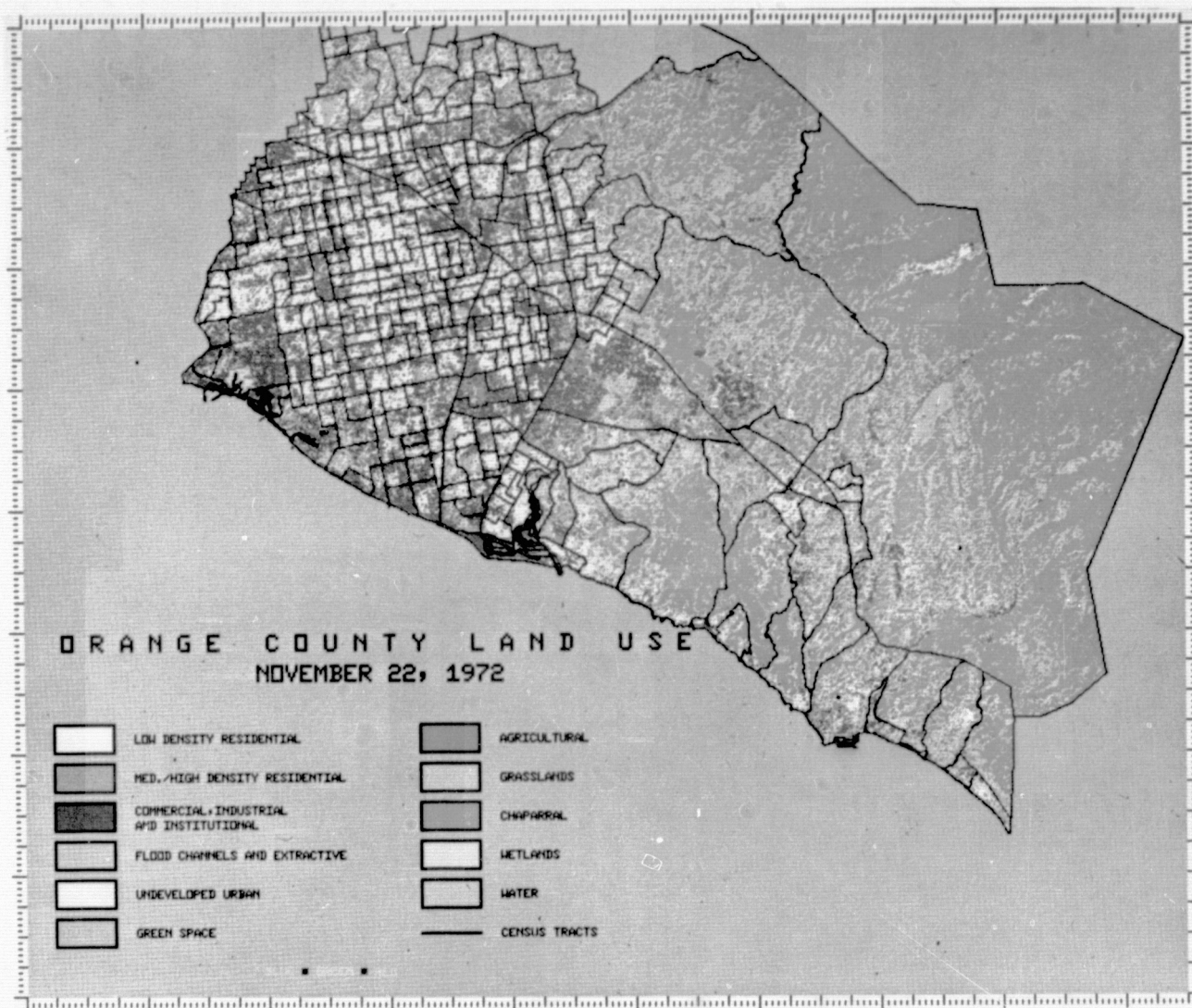
Figure 5a. Orange County Land Use Interpreted
from Multispectral Classification of
LANDSAT Digital Imagery

8. U.S. Bureau of the Census, Census Use Study: The DIME Geocoding System, Report No. 4, Washington, DC, 1970.

9. Deuker, K. J., (1972), op. cit.
Deuker, K. J., (1974), op. cit.
Werner, P. A., op. cit.

10. Dyer, H. L., et al., op. cit.
Miller, W. R., op. cit.
Power, M. A., op. cit.

11. Cooke, D., op. cit.
U.S. Bureau of the Census, op. cit.

12. Chrisman, N. R., "Topological Information Systems for Geographic Representation," presented at The Second International Symposium on Computer Assisted Cartography, Autocarto II, Reston, VA, September 22, 1975.

13. Bryant, N. A., and A. L. Zobrist, "IBIS: A Geographic Information System Based On Digital Image Processing and Image Raster Datatype", Proceedings of the LARS Symposium on Machine Processing of Remotely Sensed Data, Lafayette, Indiana, 29 June 1976, p. 1 A-1.

14. Dyer, H. L., et al., op. cit.
Power, M. A., op. cit.

15. Southern California Association of Governments, "Current and Future Land Use in SCAG Region", typescript, 15 July 1976.

16. Deuker, K. J., and F. E. Horton, "Urban-Change Detection Systems: Remote Sensing Inputs," Photogrammetria, Vol. 29 (1972), pp. 89-106.

Figure 5b. Orange County Land Use Interpreted
from Aerial Photography 1972

17. Nelefson, R., P. H. Swain, and J. R. Wray, Urban Land-Use Mapping by Machine Processing of ERTS-1 Multispectral Data: A San Francisco Bay Area Example, Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, IN, LARS Information Note 101573, 1973; Anderson, J. R., Hardy, E. E., and Roach, J. T., "A Land-Use Classification System for Use with Remote Sensor Data," U.S. Geological Survey Circular 671, Washington, D.C.

18. Addington, J. D., "A Hybrid Classifier Using the Parallelepiped and Bayesian Techniques, "Proceedings of the American Society of Photogrammetry, 41st Annual Meeting, March 1975, pp. 772-784.
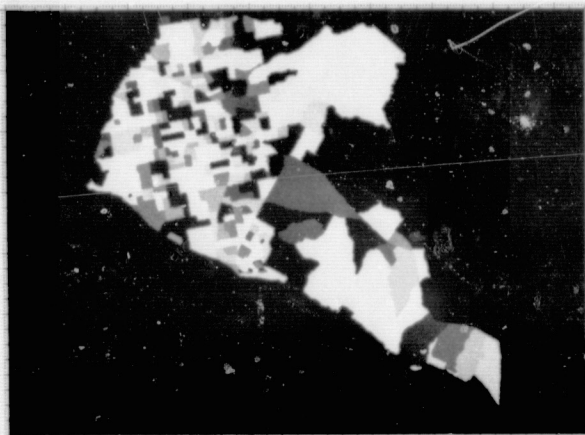
Figure 6a. Orange County Census Tract
Painted Image
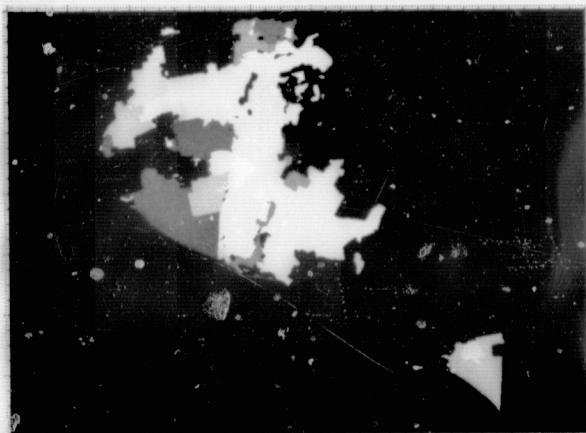
Figure 7a

Figure 6b. Orange County Municipal
Boundaries Painted Image

Figure 7b

Figure 7. Tabulations of Orange County Land
by Census Tracts and Municipalities

NEW TECHNIQUES FOR IMAGE MOTION COMPENSATION

John E. Kreznar

Jet Propulsion Laboratory
Pasadena, California

ABSTRACT

The transfer function due to spatially invariant image motion need not necessarily have zeros. This paper describes initial explorations of two techniques for eliminating these zeros. The first technique corresponds not to a physical change in the imaging process, but to an unconventional mapping between the unsampled and the sampled statements of the problem; from the new perspective thus attained, "transfer function" takes on a modified meaning wherein it remains useful for image restoration but its zeros can be eliminated. The second technique involves modulating the transmittance of the shutter during exposure in such a way that signals of all spatial frequencies are transmitted with at least some non-zero amplitude, thereby making it possible to restore them.

INTRODUCTION

The one-dimensional analog of space-invariant image motion blur may be expressed by one of the well known convolutional relations

$$q(x) = \int_{y=-\infty}^{\infty} f(x-y)p(y)\,dy + n(x) \qquad (1)$$

or

$$q_j = \sum_{k=-\infty}^{\infty} f_{j-k}p_k + n_j \qquad (2)$$

depending on whether the problem is being discussed in the unsampled or sampled case. In both (1) and (2), f is the point spread function (PSF) kernel characterizing the image motion (possibly combined with other blurring effects), p is the signal that would have been recorded in the absence of distortion, n is an additive noise term, and q is the signal that actually was recorded. The significance of the infinite limits on the integration and summation in (1) and (2) is that the boundary of the image is being deliberately swept out of sight; the emphasis is to be on the effects in the interior of the image.

As is well known, Fourier transform techniques are very useful both in the theoretical study and practical solution of this problem. The transform appropriate to Eqs. (1) and (2) are, respectively

$$A(\omega) = \int_{-\infty}^{\infty} a(x)\,e^{-ix\omega}dx \qquad (3)$$

and

$$A(e^{i\theta}) = \sum_{j=-\infty}^{\infty} a_j\,e^{-ij\theta} \qquad (4)$$

As suggested by (3) and (4), an upper case letter will generally be used in this paper to denote the transform of the signal or kernel denoted by the corresponding lower case letter. It should always be clear from context whether (3) or (4) is applicable. Note that (4) is actually a special case of the z-transform

$$A(z) = \sum a_j z^{-j} \qquad (5)$$

Using these conventions combined with the applicable convolution theorem, both (1) and (2) have the same formal expression in transform space:

$$Q = FP + N$$

The transform, F, of the PSF is called the transfer function of the system.

The _ex post facto_ image motion compensation problem is here considered to be the determination or estimation of the undistorted signal, p, given the PSF, f, and the distorted signal, q. "Determination" of p will only occur in contrived contexts where noise, n, is zero and f is nonsingular. "Estimation" is more appropriate to the real problem. However, the noiseless, nonsingular case is of interest for purposes of deepening our understanding of the problem.

Almost all studies of image motion blur of which this author is aware have concentrated on cases where the transfer function due to the image motion has zeros at various spatial frequencies (Refs. 1-5). Two exceptions are an early work by Shack (Ref. 6) wherein the effects of shutter function and non-uniform motion are considered, and a work by Bryngdahl and Lohmann (Ref. 7) which is the only previous effort of which we are aware to deliberately modulate the transmittance of the shutter for purposes of subsequent smear compensation.

TRANSFER FUNCTION ZEROS, ANNIHILATION, AND SINGULARITY

The deleterious effect of zeros in the transfer function will now be reviewed. First, we note that the most frequently described instance of image motion blur has many such zeros. In this instance, image motion is spatially invariant and of constant velocity, and the shutter is assumed to switch instantaneously between its full open and full closed states. This produces, in the one-dimensional analog of Eq. (1), a PSF which is a rectangle function. This PSF and the magnitude of its transform are shown in Figs. 12a and 12b. The actual form of the transfer function is the familiar sinc function,

$$F(\omega) = \frac{2\sin a\omega/2}{a\omega}$$

where the parameter a is the spatial extent of the motion. Thus, the transfer function has equally spaced zeros at spatial frequencies which are integer multiples of $2\pi/a$ all the way out to infinity.

Each of these zeros corresponds to a distinct input signal, namely a sinusoid of a particular spatial frequency, which is annihilated due to the effects of the PSF, f. This means that the output, q, is insensitive to the presence of these signals in the input, p, whether alone or added to other input. The use of the term "annihilation" is intended here to suggest its use in linear algebra. For example, a finite, square matrix can annihilate a non-zero vector only if the matrix is singular. Thus we shall consider PSFs, f, or their transforms, F, to be singular if F has any zeros.

It is immediately clear that the noiseless, singular case cannot be solved for p; for there are components of p, namely those annihilated by f, to which any amplitude whatever can be assigned and still be consistent with the known f and q. Conversely, the noiseless, nonsingular case may be solvable for p. The exact further conditions required for this converse to hold are beyond the scope of this paper and consist mainly of carefully defined spaces of signals, p and q, and kernels, f.

In the practical case, with noise, the implications of singularity are less clear-cut. This is because in this case additional a priori statistical or probabilistic information about p and n is combined, explicitly or implicitly, with knowledge of q and f in the process of estimating p. The additional information could, for example, allow even the annihilated components of p to be estimated. It is usually the case, though, that singularity is as deleterious in this case as in the noiseless case.

Another important practical consideration is that of near-singularity. A simple but useful measure of nearness to singularity is the ratio of the maximum to the minimum of the square-magnitude of the transfer function. [In the unsampled case of Eq. (1), only a limited frequency range can be considered for this maximum and minimum since the transfer function usually goes to zero at infinite frequency.] Note that this ratio goes to infinity in singular cases. This ratio is closely related to the condition number of a matrix which is a useful measure of the difficulty of computing the inverse of that matrix (Ref. 8). In this paper we shall use "condition number" to refer to the ratio described above. Note that the condition number can also be used as a rough indication of the noise sensitivity of the restoration process, simply because the transfer function minimum is likely to be the point in the spectrum where the signal-to-noise ratio is worst.

Referring again to the space of square matrices of some finite size, we are impressed by the fact that almost all of them are nonsingular and that all those which are singular have a near neighbor which is nonsingular. Why then the apparent ubiquitousness of singular PSFs, especially in the sampled case of Eq. (2), which is so matrix-like? The question that motivates the work presented here is: How real and how tenacious are the zeros and near-zeros arising in the image motion compensation problem? Can steps be taken to eliminate zeros and decrease the condition number?

We present here initial exploration of two techniques to accomplish just this.

ALTERNATE FORMULATIONS OF THE SAMPLING PROCESS

In this technique, we assume that only the samples of the smeared image are available and no opportunity exists for more refined measurements of an original unsampled record of the image. (This is true, for example, of all recent spaceborne imaging systems.) Furthermore, we arbitrarily define the problem to be the reconstruction of samples that might have resulted had all conditions been identical except for image motion. This forces upon us the task of expressing in the sampled form of Eq. (2) the phenomenon of image motion, which is essentially unsampled and continuous. It can readily be demonstrated that this cannot, in general, be done. The basic problem is that the space of unsampled images is so rich that it is easy to find a pair of unsampled images whose smeared samples are identical but whose unsmeared samples differ and another pair whose unsmeared samples are identical but whose smeared samples differ. For later reference, we call this the mapping problem between the unsampled and sampled cases.

The mapping problem forces us to somehow limit the space of admissible unsampled images and agree to simply ignore the effects of components of the real image which fall outside this space. An often used, well known space which is free of the mapping problem is that of Nyquist bandlimited images. Our approach here is to consider alternative spaces, also free of the mapping problem, for which the consequences of ignoring components outside the space are acceptable, but for which singularities are eliminated and the condition number is driven down.

The selection of alternate spaces does not, of course, in any way reduce the actual degradation of the signal. It does, however, lead to alternate restoration filters which may produce superior results and/or be more economical of computer time.

There is a class of spaces of unsampled signals, each of which is free of the mapping problem, which might be called the "pulse-composed" signals in that each signal consists of the additive composition of pulses, one at each sample site. The particular "shape" of the pulse (its functional definition) defines the space within the class and is fixed from sample site to sample site within one such space; the various signals within the space are the result of the amplitudes of the pulses at each site.

Reasonable pulse functions will have properties such as going rapidly to zero as the distance from the center grows and having a characteristic "diameter" on the order of the sample interval. Very often the pulse will be symmetrical.

For simplicity in the present discussion, we arbitrarily restrict our attention to pulses $s(x)$, for which $s(j-k) = \delta_{jk}$. This says simply that the pulse has unit value at its center and zero value at all non-zero integer values of its argument. We also assume that a sample of a signal is precisely its value at the sample site; this effectively means that any non-zero spatial extent of

the physical sampler's sensitivity has already been convolved with the signals. These simplifying assumptions enable us to write the following relationships between the unsampled and the sampled signals:

$$p(x) = \sum_j p_j s(x-j) \qquad (6)$$

$$p_j = p(j)$$
$$q_j = q(j) \qquad (7)$$

In the context just described, it becomes a simple matter to write an expression of sampled form (2) which is equivalent to the unsampled Eq. (1). Before doing so, however, we rewrite (1) in the form

$$q(x) = \int_y f(x-\xi-y)\, p(y)\, dy \qquad (8)$$

wherein the noise term has been dropped as being irrelevant to the present discussion, and wherein an offset parameter, ξ, has been incorporated. In the unsampled context of (8), this offset has the simple effect of translating signal q by displacement ξ. This offset plays a significant role in the sampled case, however, and this happens to be the simplest way to introduce it into the mathematics.

Finally, the following development yields the desired sample convolution. Substituting (6) into (8),

$$q(x) = \int_y f(x-\xi-y) \sum_j p_j s(y-j)\, dy$$

$$= \sum_k p_k \int_y f(x-\xi-y) s(y-k)\, dy$$

$$= \sum_k p_k \int_y f(x-k-y-\xi) s(y)\, dy$$

Substituting this result into (7),

$$q_j = \sum_k p_k \int_y f(j-k-y-\xi) s(y)\, dy$$

or

$$q_j = \sum_j f_{j-k} p_k \qquad (9)$$

where

$$f_j = \int_x f(j-x-\xi) s(x)\, dx \qquad (10)$$

This demonstrates that a space of pulse-composed functions is free of the mapping problem and in fact explicitly displays the mapping from the unsampled to the sampled formulation.

It may be noted at this point that the Nyquist bandlimited mapping is a special case of the foregoing which results upon substituting $s(x) = (\sin \pi x)/(\pi x)$.

One-Dimensional Case

As a supplement to the two-dimensional results to follow, a very cursory one-dimensional study has been performed based on the development given above. The pulse-function which was used is a simple triangle

$$s(x) = \begin{cases} 1-|x| \,, \ |x| < 1; \\ 0 \,, \ |x| \geqslant 1. \end{cases}$$

The smear-functions, $f(x)$, were simple rectangles of various lengths not less than two samples, which is the onset of singularity in the Nyquist bandlimited case. Figure 1 shows an array of the pulse-functions on one sample centers and graphically suggests the integration given in Eq. (10).

What was done was to locate the roots of the z-transform (5) for various values of the smear length and of the offset, ξ. The results may be summarized as follows.

If the smear is of integer length not less than 2, the weights f_j are of the form

$$a,b,1,1,\ldots,1,1-a,1-b$$

where the number of 1's is zero or more and a and b depend on the offset, ξ. A simple calculation shows that $\exp(2\pi i)/(n-2)$ is a root of the z-transform, where n is the total number of non-zero weights; this root being on the unit circle means that in all these cases, the problem is singular.

The situation improves somewhat for some non-integer values of smear, although in this case, the only information available is numerical root computations; no general result concerning singularity has been obtained. Based on a few numerical results at smear lengths of 2.2, 3.1, 3.2, 3.5, 6.1, and 6.5 samples, it appears that a) a one-to-one correspondence can be established between the unit circle zeros in the Nyquist bandlimited case and those in the $s = 1-|x|$ case; b) as a function of offset ξ, the roots wander in the vicinity of the corresponding Nyquist roots, usually off the unit circle; c) there is evidence that the greatest distance between the unit circle and the nearest root thereto which can be achieved by manipulating offset ξ falls off rapidly with increasing smear length. When the smear length was 2.2 samples, the root nearest the unit circle could be moved out to magnitude > 1.49 at offset ξ of .2. But at 6.5 samples smear, the largest departure from the unit circle of the worst case root, the one at $\theta \approx 55.5°$, had a magnitude of 1.001136 at $\xi = .001$. This is so discouragingly close to the unit circle that computations of the condition number for these cases were never performed; it surely would have come out rather large.

Two-Dimensional Case

The generalization to two dimensions of the technique just described has been developed into a

complete image motion compensation computer program. This program operates as follows. The weights of a sampled kernel are computed by a two-dimensional analog of the scheme above. A two-dimensional finite discrete Fourier transform (DFT) is then employed to evaluate the z-transform of this kernel on a fine mesh of points on the product-set of the two unit circles (the "unit torus"). Manipulations are performed upon this transform whose scope will eventually comprehend the simple inversion, Wiener filtering, and constrained least squares techniques (Ref. 9), but not all of which have been implemented at this writing. A two-dimensional DFT is again employed at this point to approximate the inverse z-transform, yielding the weights of an image motion compensating kernel, or filter. This filter is then applied, using direct convolution, to the smeared image. (The images involved have exceeded 1024x1024 samples so it has been impractical to do the filtering in Fourier space.)

The program can optionally synthesize a checkerboard image of the compensating kernel wherein each square corresponds to one element of the kernel and has a gray level representing its value. Some examples are shown in Fig. 2. All four of the filters illustrated in this figure are designed to compensate a 4.0-pixel smear 17° from horizontal. To explain why they are different, we now amplify on that aspect of the program which is pertinent to the subject of this paper.

Using a development analogous to that used in one dimension, it can be shown that, if the unsampled incident signal is pulse-composed, that is, if it is of the form

$$p(x,y) = \sum_j \sum_k p_{jk} s(x-j,y-k)$$

where $s(x,y)$ is a suitable two-dimensional pulse function, then for any unsampled PSF, $f(x,y)$, there exists a sampled kernel $\{f_{jk}\}$ such that if

$$q(x,y) = \int_u \int_v f(x-u,y-v)p(u,v)\, dv\, du$$

then

$$q_{jk} = \sum_m \sum_n f_{j-m,k-n} p_{mn}$$

where $\{q_{jk}\}$ and $\{p_{jk}\}$ are the samples of $q(x,y)$ and $p(x,y)$. As in the one-dimensional case, an arbitrary offset enters into the computation of $\{f_{jk}\}$. A sketch suggesting the relationship between the unsampled and the sampled kernels is given in Fig. 3. In this figure, each of the "knobs" on the regular array corresponds to a copy of the pulse function, $s(x,y)$, situated at a sample site. The unsampled kernel is repeated twice to illustrate two different possibilities for the offset. It can be thought of as a rectangular solid of infinitesimal thickness, width proportional to the smear length, and oriented in the direction of the image motion. This figure merely illustrates the integrals

$$f_{jk} = \int_x \int_y f(j-x-\xi,k-y-\eta)s(x,y)\, dy\, dx$$

which yield the correct sampled kernel weights, f_{jk}.

The difference among the filters displayed in Fig. 2 can now be explained. They are due solely to variations in the offset parameters, ξ and η. Figure 2c is a dramatic illustration of one of the real benefits of decreasing the condition number, namely a large reduction in the number of significantly non-zero terms in the compensating filter. Though their results may be comparable in quality, the filter of Fig. 2c involves far fewer multiply-adds than that of Fig. 2a. The difference easily can amount to an hour of computer time for one picture.

Concluding this description of this technique for singularity avoidance, it may be said that no practical method is yet known to optimize the choice of the offset, (ξ,η). Nor has work been extended to other than the simple triangle pulse function and its two-dimensional analog. In short, the full potential of this technique is unknown.

THE USE OF SHUTTER MODULATION

The shutter function is defined as the shutter transmittance as a function of time. The PSF resulting from uniform image motion will thus be modulated by the shutter function. If a camera is equipped with a shutter for which this function can be other than a simple rectangle, then intelligent selection of this function can be used to eliminate singularities. This shutter modulation technique may be regarded as a generalization of that of Bryngdahl and Lohman (Ref. 7) alluded to in the Introduction.

Throughout the following it is well to bear in mind the constraints placed upon the shutter function. First, the speed of the lens defines a certain maximum which can never be exceeded. Second, the shutter function can never go negative. Finally, the area under the shutter function, which is proportional to the total light admitted, must provide adequate S/N ratio in the recorded image. In this third constraint, "adequate S/N ratio" is deliberately left vague. In fact, a tradeoff exists between S/N ratio and degree of smear, the details of which are beyond the scope of this paper.

The simple rectangular PSF can be made an even function by centering it over the origin. It is not difficult to perceive why zeros are commonplace in the transfer functions of even PSFs nor to perceive a simple scheme for eliminating them by adding an odd perturbation to the even PSF. The Fourier transform of an even function is real; therefore, if there is any phase reversal at all, the transform must pass through zero. As soon as an odd component is allowed, however, the Fourier transform as a function of frequency is no longer confined to the real axis, but is free to wander elsewhere in the complex plane. Phase reversals are much more likely to occur by spiraling around complex zero than by passing right through it. Indeed, the odd component would have to be very carefully chosen in order to preserve the zeros.

The shutter modulation technique involves nothing more than exploiting this property of odd functions in order to not only eliminate the zeros of the transform function, but to maximize the "miss-distance" from zero in the complex plane.

Notice that the technique achieves its goal of eliminating singularity and reducing condition number at the expense of symmetry and simplicity. The phase behavior of the transfer function becomes especially messy and must be properly compensated in the restoration process (Ref. 10).

The Sampled Case

A few very simple cases were selected in the sampled case just to get a rough idea of the potential of the technique. They are illustrated in Figs. 4 through 11. The "a" parts of these figures show the two or three non-zero elements of the kernel $\{f_j\}$, and the "b" parts show the magnitude of the transfer function over its entire range from $-\pi$ to π. In all cases the sum of the elements of $\{f_j\}$ is the same and may be taken to be unity.

Figures 4 through 6 illustrate the singular behavior generally found for equal-weight kernels. Though equal-weight kernels are an attractive, simple means to gather enough light subject to an imposed maximum upon the individual weights, they should clearly be avoided because of this singularity.

No weight used in Figs. 4 through 8 exceeds 1/2, while Figs. 9 through 11 illustrate the result of allowing weights up to 3/4 due to a faster lens. Of the cases where the weights are limited to 1/2, only one, Fig. 8, is non-singular, clearly demonstrating the virtue of asymmetric distribution of the weight. The weights used in this case are 1/6, 1/3, and 1/2. This turns out to be the global optimum among 3-weight cases limited to a maximum weight of 1/2; the condition number is 13.5.

In Figs. 9 and 10 the weight 3/4 element was placed at the center and only the distribution of the remaining 1/4 between left and right changes; both cases are nonsingular and have a condition number of 4. (In fact, the condition number is independent of how this remaining 1/4 is distributed between left and right.) This approach of putting the maximum weight at the center of three elements is a local optimum in the sense that any perturbation which removes a bit of weight from the center element and places it on the left and/or right will enlarge the condition number. The lowest maximum weight for which this local minimum fails to occur is 1/2; in other words, Fig. 7 illustrates a case which is just singular, with any slight migration of weight to the center element leading to a nonsingular case.

Figure 11, of course, is the globally optimum case which supports the general conclusion that highly asymmetric cases are often the best. In this case the weights are 1/12, 1/6, and 3/4, and the condition number is only $81/32 \approx 2.53$.

The Unsampled Case

Estimating the potential benefits of shutter modulation in the unsampled case is more difficult than in the sampled case because the space of possible shutter functions is so rich and because no simple, natural performance measure is available. In order to get some crude idea of what is possible, we considered two families of shutter functions, trapezoids and pairs of rectangles. The interest in the latter arises just because it may shed some light on what can be done with a conventional "bang-bang" shutter if multiple shutterings are permitted.

The results of selecting various trapezoidal and rectangle-pair shutter functions are shown in Figs. 12 through 19. In all cases, the integrated transmittance is the same and the maximum value of the function is the same. The transfer function is shown for frequencies out to approximately the second zero of the "reference" case, the rectangle of Fig. 12.

The generally inferior performance of symmetric shutter functions is evident. Also evident is the definite advantage of multiple shutterings when the shutter must be switched instantaneously between its full closed and full opened states.

CONCLUSIONS

Two techniques for reducing singularity and near-singularity of the transfer function due to image motion have been examined. The first, alternate formulations of the sampling process, is merely a conceptual departure leading to a new perspective on the problem and has been embodied in an image motion compensation computer program now in active development. The second, shutter modulation, corresponds to a real change in the imaging process. Both have been shown to be capable of not only eliminating the zeros in the transfer functions, but increasing the smallest value taken on by the transfer function. Since this smallest value is an accurate measure of the true degradation of the image, any increase of it makes the image more effectively restorable. Much work remains, however, to determine the full potential of the techniques.

REFERENCES

1. Mazurowski, M. J., and Kinzly, R. E., "The Precision of Edge Analysis Applied to the Evaluation of Motion-Degraded Images," NASA SP-193, Evaluation of Motion-Degraded Images, Office of Technology Utilization, National Aeronautics and Space Administration, Washington, DC, 1969.

2. Harris, J. L., Sr., "Potential and Limitations of Techniques for Processing Linear Motion-Degraded Imagery," ibid.

3. Cutrona, L. J., and Hall, W. D., "Some Considerations in Post-Facto Blur Removal," ibid.

4. Roetling, P. G., Haas, R. C., and Kinzly, R.E., "Some Practical Aspects of Measurement and Restoration of Motion-Degraded Images," ibid.

5. Sawchuk, A. A., "Space-Variant Image Restoration by Coordinate Transformations," J. Opt. Soc. Am. Vol. 64, No. 2, p. 138, February 1974.

6. Shack, R. V., "The Influence of Image Motion and Shutter Operation on the Photographic Transfer Function," Appl. Opt., Vol. 3, No. 10, p. 1171, October 1964.

7. Bryngdahl, O., and Lohmann, A., "Compensation of Motion Blur by Shutter Modulation," NASA SP-193, Evaluation of Motion-Degraded Images, Office of Technology Utilization, National Aeronautics and Space Administration, Washington, DC, 1969.

8. Ekstrom, M. P., "A Spectral Characterization of the Ill-Conditioning in Numerical Deconvolution," IEEE Trans. Audio Electroacoust., Vol. AU-21, No. 4, p. 344, August 1973.

9. Hunt, B. R., "Digital Image Processing," Proc. IEEE, Vol. 63, No. 4, p. 693, April 1975.

10. Huang, T. S., Burnett, J. W., and Deczky, A.G., "The Importance of Phase in Image Processing Filters," IEEE Trans. Acous. Speech and Sig. Proc., Vol. ASSP-23, No. 6, p. 529, December 1975.

Figure 1. Visualization of the integrals relating the unsampled and the sampled kernels in one dimension.

2a
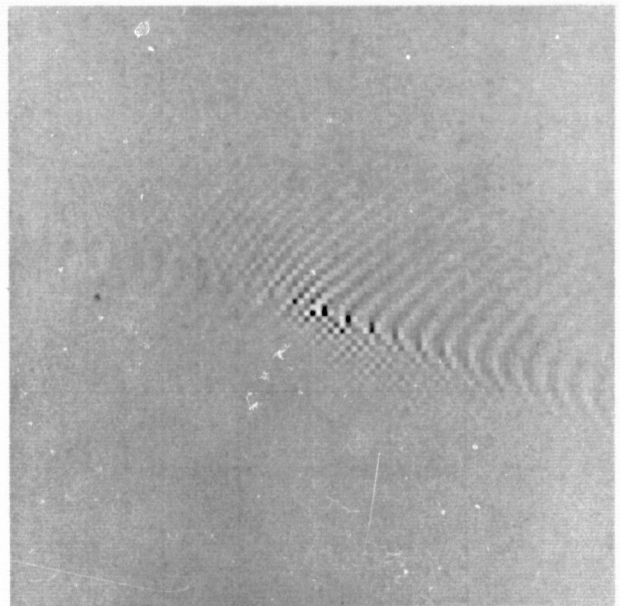
2b
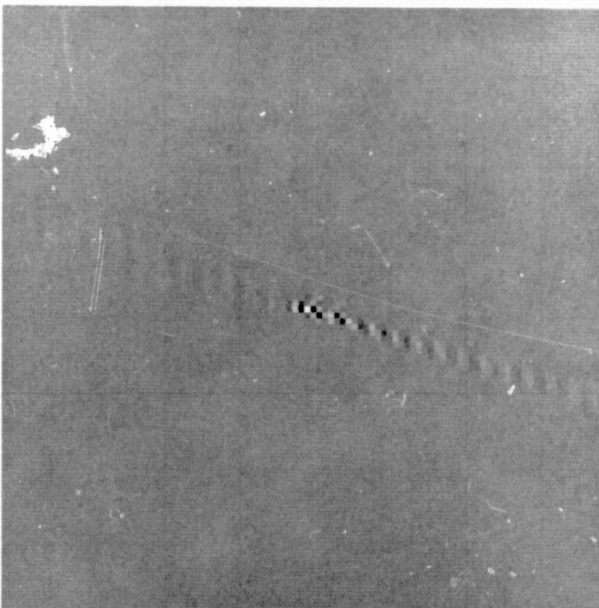
2c
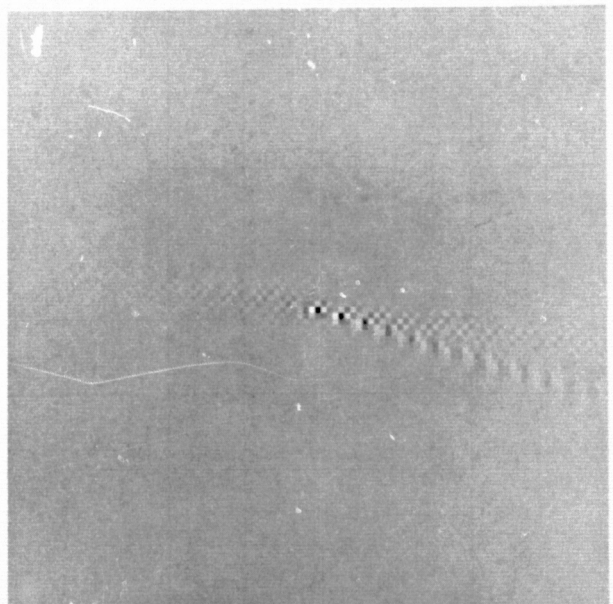
2d

Figure 2. Gray-level representation of several two-dimensional image
motion compensation filters.
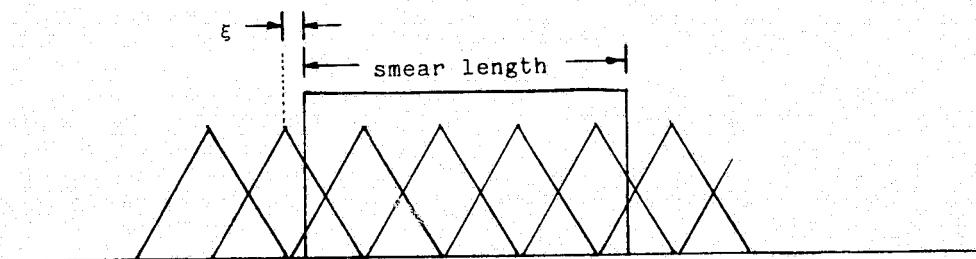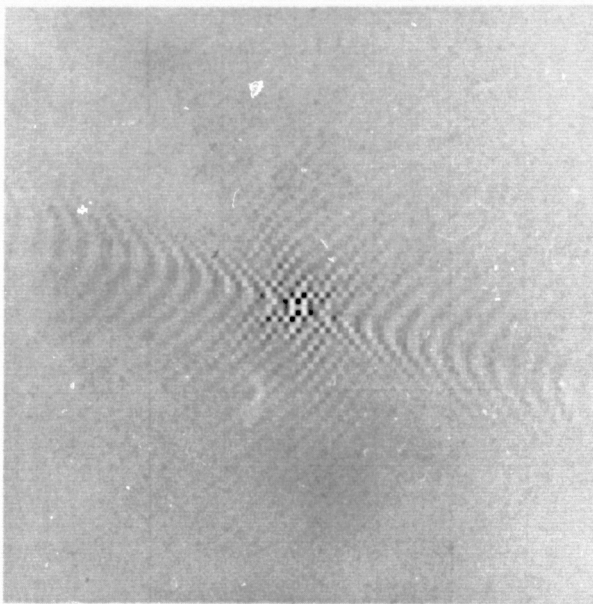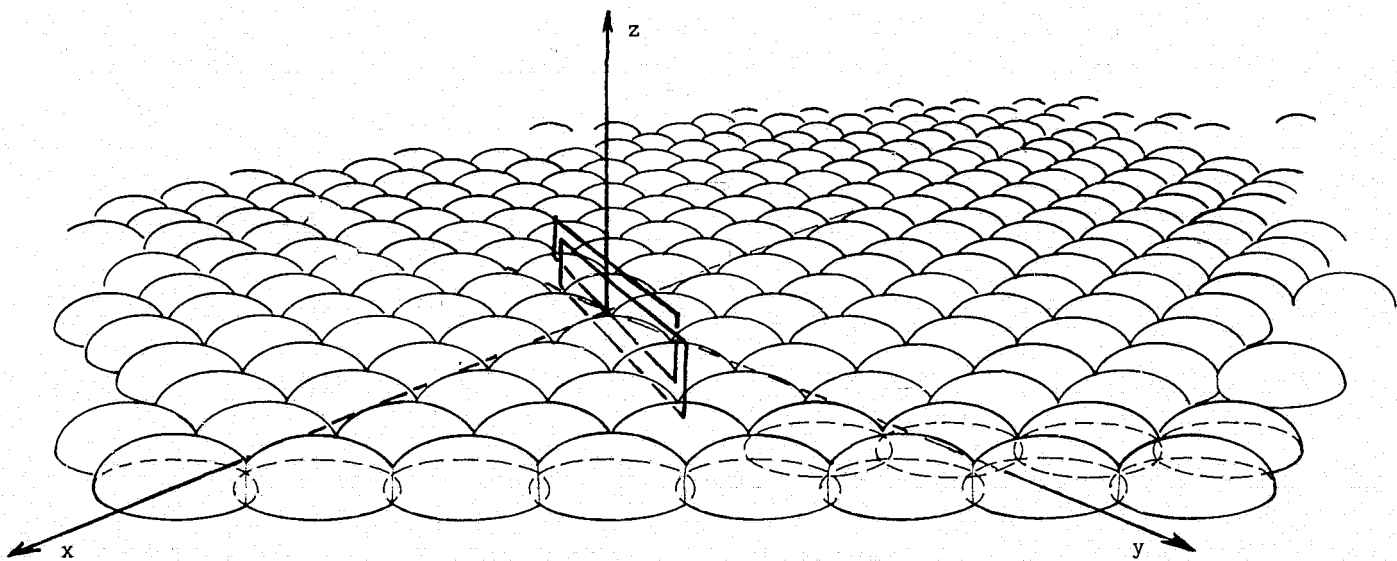
Figure 3. Visualization of the integrals relating the unsampled and the sampled kernels in two dimensions.

4a

4b

5a

5b

6a

6b

7a

7b

8a

8b

9a

9b

10a

10b

11a

11b

12a

12b

13a

13b

14a

14b

15a

15b

16a

16b

17a

17b

18a

18b

19a

19b
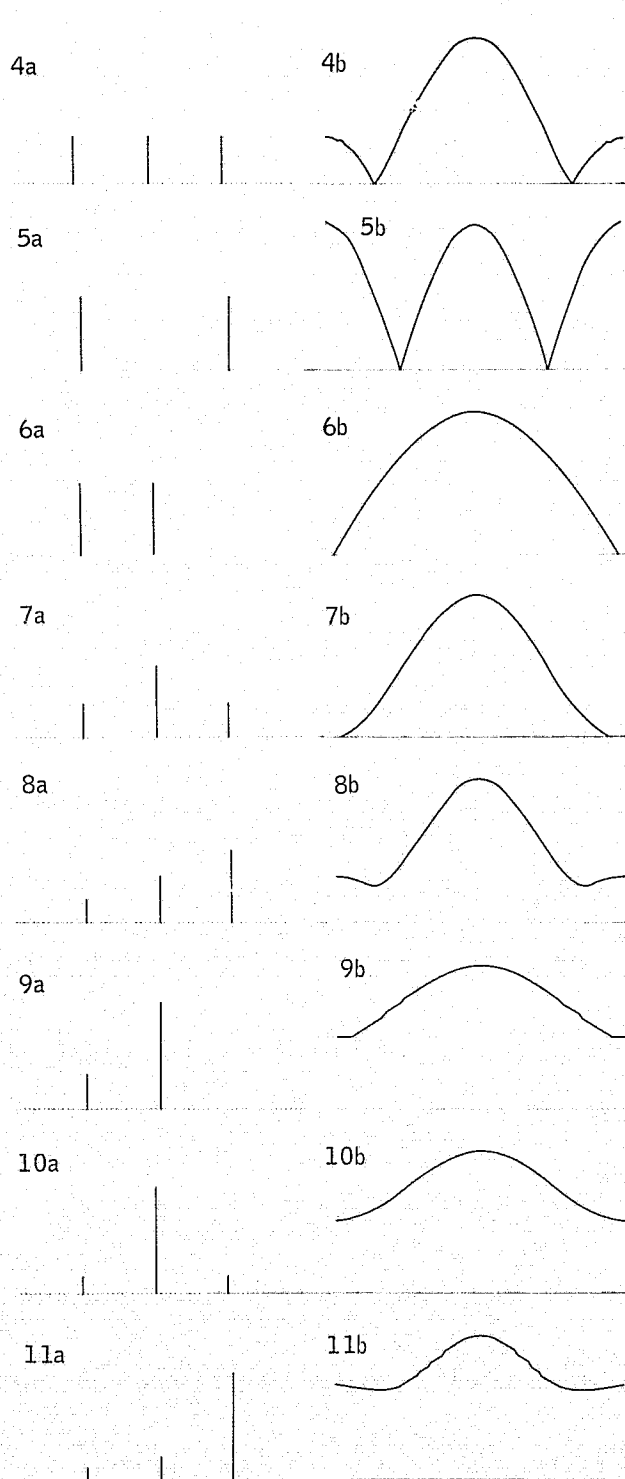
Figures 4-11. Shutter-modulation examples in the sampled case.
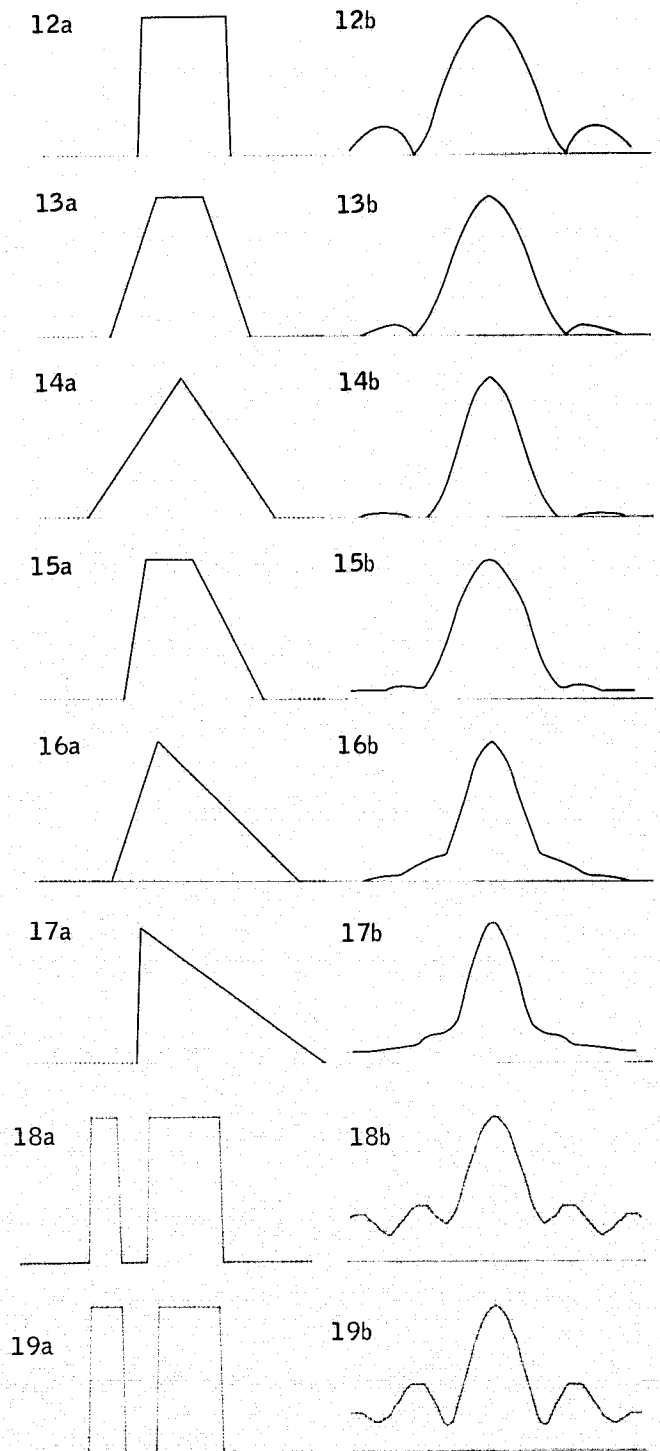
Figures 12-19. Shutter-modulation examples in the unsampled case.

DERIVATION OF A THERMAL INERTIA IMAGE
FROM REMOTELY SENSED DATA

Anne B. Kahle

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

ABSTRACT

Thermal inertia of the Earth's surface can be used in geologic mapping as a complement to surface reflectance data as provided by Landsat. Thermal inertia cannot be determined directly but must be inferred from radiation temperature measurements (by thermal IR sensors) made at various times in the diurnal cycle, combined with a model of the surface heating processes. We have developed a model which differs from those created previously for this purpose because it includes the sensible and latent heating. Tests of this model using field data indicate that it accurately determines the surface heating. When the model is used with field measurements of meteorological variables, and combined with remotely sensed temperature data, a thermal inertia image can be produced.

I. INTRODUCTION

As part of our continuing program in the application of remote sensing to mineral exploration, we are examining the utility of data from all parts of the electromagnetic spectrum which are amenable to remote sensing. Thermal (infrared) emission from the Earth can be sensed remotely in the atmospheric windows, particularly between 8 and 14 μm. The measured radiation can be related simply to the temperature of the surface. From measurements of the diurnal variations in temperature combined with a consideration of the processes acting to heat and cool the surface, it is possible to infer the thermal inertia of the surface material (Ref. 1, 2, 3). This thermal property of the surface material can be used in lithologic mapping as a complement to other remotely sensed data. Thermal inertia is a volume property rather than a surface property. Thus a knowledge of thermal inertia permits discrimination among materials which may be indistinguishable when only surface reflectivity data is available. This remote sensing of a volume property is also particularly useful in geologic applications because exposed surfaces are often weathered, stained, or covered with a thin layer of another material. Thermal inertia may also be of use in distinguishing outcrop from alluvium.

The derivation of a thermal inertia image from remotely sensed data requires extensive use of the computer, both for surface heat balance modeling and for image processing. This paper describes some of the techniques we have developed to solve this problem. The basic idea is to first calculate theoretically the heating which occurs at the site. Next, a model is used to determine theoretically the variation of the temperature of surface material through the diurnal cycle in response to this heating, for a range of possible values of thermal inertia. Then the actual surface temperature is measured, using a line scanner, at times of maximum and minimum temperature. Several image processing steps are applied to these measured data

sets to relate the measured diurnal temperature variations to the theoretically determined values. Finally, a thermal inertia image is created, based on this relationship.

II. THERMAL MODEL

A surface heating model has been developed from which we compute temperature of the surface material as a function of depth and time. The one-dimensional heat conduction equation

$$\frac{\partial T}{\partial t} = \kappa \frac{\partial^2 T}{\partial z^2} \tag{1}$$

where
T = temperature

t = time

κ = thermal diffusivity = $\frac{K}{\rho C}$

z = depth

K = thermal conductivity

ρ = density

C = specific heat

is solved numerically by difference equations, for the top 50 cm of the surface. The boundary conditions are that the heat flux, G, across the lower boundary, $z = 50$ cm, is zero, and that the heat flux across the upper boundary, $z = 0$ cm, satisfies the heat balance

$$S + R + H + L + G = 0 \tag{2}$$

where S = net solar radiation

R = net thermal radiation from the atmosphere and the ground

H = sensible heat flux between the atmosphere and the ground

L = latent heat flux between the atmosphere and the ground

G = heat flux in the soil

All fluxes are considered to be positive if directed toward the earth-atmosphere boundary. Determination of these boundary heat flux terms constitutes a major part of the modeling problem. Details of the expressions we use for these terms are given by Kahle (Ref. 4).

We use an expression for the solar radiation developed by Joseph (Ref. 5, 6) and Arakawa et al. (Ref. 7), which is a function of time of day, time of year, water vapor content of the atmosphere, ground albedo, atmospheric scattering albedo, cloudiness, and site elevation. Most of these can be estimated or calculated without on-site measurements. Measurement of the ground albedo, which is essential, can be done remotely. An important

modification to the solar radiation calculation which must be made in most cases is the inclusion of topography, because the slope and orientation of the surface significantly affect the amount of radiation received by the surface.

The long-wave radiation from the ground depends on the ground temperature and emissivity which, for simplicity, we assume to be unity. This assumption has been investigated and appears to have only a second order effect on the thermal inertia calculation because we measure day-night temperature differences rather than absolute values of the temperature. Computation of the ground temperature is described below. The downward long-wave radiation from the atmosphere is a function of the amount of water vapor and other emitters, and the temperature of the emitting regions. This can be expressed in terms of radiation at an "effective sky temperature". We use typical values of this sky temperature given by Kondratyev (Ref. 8).

The expressions we use for the sensible heat and latent heat transfer, H and L, were originally developed by Burke (Ref. 9). In addition to being a function of ground temperature, these terms depend upon the value of the air temperature, wind speed and humidity at a single, representative level near the ground, so on-the-site measurements of these quantities are required. We are investigating the feasibility of determining these by interpolation from neighboring National Weather Service stations, modified by knowledge of the local meteorological phenomena.

The final term in the surface heat balance condition is the heat flux into the ground $G = K(\partial T/\partial z)$. In our difference equation formulation, this expression contains the temperature at the surface and at the first level below the surface, and also the thermal conductivity, K, which is the thermal property we are seeking to determine.

The complicated and non-linear boundary condition (eqn. 2) requires that we numerically solve the heat conduction equation (eqn. 1). A finite difference formulation is used. We assume the surface material is homogeneous down through the depth to which the diurnal heating wave penetrates. The lower boundary is chosen to be a depth z = 50 cm. This surface layer is subdivided into one cm thick layers. At an initial time, t = 0, the temperature is defined for each layer. The temperature at the next time step, t = 1, is then computed for each layer from the finite difference form of the heat conduction equation,

$$T(t + 1, z) = T(t, z)$$
$$+ \kappa \frac{\Delta t}{\Delta z^2} \left[T(t, z + 1) - 2T(t, z) + T(t, z - 1) \right]. \tag{3}$$

The updated value of the surface temperature $T(t + 1, 0)$ is calculated by solving for the temperature, T_g, which will satisfy the surface heat balance equation (eqn. 2) which can be written in the form

$$T_g^4 + C_1 T_g + C_2 = 0 \tag{4}$$

This equation can be solved numerically for T_g by use of the Newton-Raphson iteration technique.

The time steps must be small ($\sim$ 50 sec) to insure stability, so satisfactory convergence is achieved after only one or two iterations.

Having thus computed the new value of the surface temperature, at t = 1, the subsurface temperatures are computed for the next time step, t = 2, again using equation 3. This procedure is repeated throughout the diurnal cycle. The final product is the temperature distribution in the ground as a function of depth and time, for a given set of values of site parameters, including meteorological conditions, topography, and thermal properties of the surface layer.

For a given set of environmental parameters, the model will produce the diurnal variation in the surface temperature. However, in the geological remote sensing problem, one measures the diurnal temperature range of the surface, ΔT, and seeks to determine the surface thermal properties. We accomplish the inversion of this solution by 1) running the model for a sufficiently complete range of values of the thermal properties, 2) creating a look-up table which gives the diurnal surface temperature range, ΔT, as a function of thermal inertia, and then 3) inverting this table to give thermal inertia as a function of ΔT.

The test of the model was conducted by making measurements of the ground surface temperature over a diurnal cycle, with a Barnes PRT-5 radiometer, and comparing these temperatures with those predicted by the model for the measured site conditions. The results of this test are shown in Fig. 1 for several sites on Lavic Lake, a playa in the Mojave Desert. It is clear that for these test sites in this particular case the model predicts the temperature variation quite accurately. The wind and air temperatures measured on the playa during the 48 hour period prior to and during the ground surface temperature measurements are shown in Fig. 2. The effects of the stronger winds during the first half of the experiment can be clearly seen in the much lower surface temperatures in the first daytime period shown in Fig. 1 and Fig. 2. The ability of the model to treat this cooling is apparent. The magnitude of each of the heating terms of equation (2) as a function of time as computed by the model is shown in Fig. 3. The latent heat flux was ignored for this test. The heating terms compare favorably with the same terms as calculated by Sasamori (Ref. 10) for dry soils and as measured by Vehrencamp (Ref. 11) on a dry lake at El Mirage, California. This version of the model was therefore utilized to determine the thermal inertia of this test site from remotely sensed data, as described in the next section, with additional details given by Gillespie and Kahle (Ref. 12).

III. CONSTRUCTION OF A THERMAL INERTIA IMAGE

To test the utility of thermal inertia for lithologic mapping, a study was undertaken at the Pisgah Crater - Lavic Lake test site. The eleven-channel Modular Multispectral Scanner (M^2S) was flown over the site in the NASA NP-3A aircraft. This scanner has 10 channels which cover the visible and near IR spectrum (.4 to 1.1 µm) and one channel in the thermal IR (8-14 µm). Four series of flights at successive pre-dawn and post-noon times were made on March 29 and 30, 1975. The

elevation of the aircraft at approximately 1200 m and 4000 m provided ground resolutions of approximately 3 m to 10 m, respectively.

Concurrent ground measurements were made of soil moisture, ground temperature with both radiometers and probes, surface albedo, and meteorological conditions. The wind and air temperature values measured during the 48 hour period prior to and during the aircraft flights were shown in Fig. 2 of the previous section. Measurements indicated that the relative humidity and ground moisture were low enough that the latent heat flux term could be ignored in the modeling.

The Jet Propulsion Laboratory's Image Processing Laboratory facilities were used to analyze the M^2S data, which were supplied to us in the form of computer compatible digital tapes by Johnson Space Center. Data from two pre-dawn and afternoon low-level flight lines on March 30 were selected for analysis. An albedo image was computed from the 10 channels of visible and near IR data, calibrated by use of the ground measurements made with the JPL Portable Field Reflectance Spectrometer in the 0.4-2.5 μm range and weighted by the solar illumination curve. The albedo on flight line 2 is shown in Fig. 4, with areas of high albedo being bright. At the top of the image are alluvial fans. In the right center, is Lavic Lake, a dry playa. At left center is the end of the Pisgah lava flow. Below Lavic Lake is the Sunshine lava flow, with Sunshine Crater at the bottom of the image.

Daytime and nighttime temperature images were created from the thermal IR radiance data, which were calibrated with on-board black-body sources. These temperature images are shown in Figs. 5 and 6. Registration of the day and night data required several steps. The first was a simple scan rectification to correct for the viewing geometry. Next, each data set was fitted to a 7-1/2 minute topographic map by GEOM, a program using a bi-linear interpolation resampling algorithm. This removed the gross effects of the varying altitude and direction of the aircraft flight line. Finally several iterations were required to register locally the two data sets to remove the effects induced by turbulence on the aircraft flight. A day-night temperature difference image, ΔT, was then generated. This is shown in Fig. 7 in which bright indicates a large temperature difference. The irregular boundaries are the result of the geometric corrections.

The JPL thermal model was then applied to this data set (albedo and ΔT) to compute the thermal inertia, ignoring topography for this calculation. This thermal inertia is shown in image form in Fig. 8, with high thermal inertia being bright. The computed values of thermal inertia centered around 0.02 cal cm^{-2} °K^{-1} sec$^{-1/2}$ for the playa and around 0.05 cal cm^{-2} °K^{-1} sec$^{-1/2}$ for basalt. These values are reasonable for these materials. It is obvious, though, that no lithologic identification could be made strictly on the basis of the thermal inertia image.

The topography was not included in the construction of the thermal inertia image in Fig. 8. To investigate the importance of topographic effects we computed the thermal inertia in the immediate vicinity of Pisgah Crater (on the other flight line) with and without the inclusion of digital topography. A first attempt was made using the commercially available 1:250,000 NCIC digital altitude data. This image did not provide sufficient resolution at the scale of the M^2S imagery so we used digital topography derived from aerial photography.

A digital image describing scene altitude was constructed in the form of a square grid, digital terrain model (DTM), with a sample interval equal to that of the M^2S image, from aerial photographs taken simultaneously with the scanner data, at a scale of 1:12,000. The elevations were sampled along parallel, equidistant profiles and recorded in digital form using standard stereo-photogrammetric procedures.

The DTM was constructed from the discreet, digital profiles using a simple and fast point-wise interpolation algorithm particularly well suited to profile data: at each pixel, a height was interpolated using the 16 closest available data points, which defined a polynomial of 3rd order. Fig. 9a shows computed insolation at Pisgah Crater based on the altitude image, assuming a Lambert surface of constant albedo.

From the digital altitude image we derived two images describing slope and slope azimuth. These values were computed at each pixel by determining the altitude gradient from the four nearest pixels. We therefore had four registered images (albedo, ΔT, slope, and slope azimuth), at Pisgah Crater. The thermal inertia was determined from these data sets by application of the model. The two thermal inertia images of this area, without and with topographic corrections, are shown in Figs. 9b and 9c. The difference in thermal inertia is shown in 9d. These clearly demonstrate the necessity for such corrections in regions of moderate to higher relief. In the uncorrected version (Fig. 9b), the crater appears anomalously bright on the east-facing side and dark on the west-facing side. This effect occurs because the east side was already shaded when the afternoon images were acquired, while the west side was fully illuminated. Thus the east side was cooler than horizontal surfaces of the same substance (basalt cinders), while the west side was warmer, and these anomalous temperatures produced the anomalous thermal inertias. In the topographically compensated version (Fig. 9c) these effects are reduced. The difference picture (Fig. 9d) clearly illustrates how the change is related to solar illumination (Fig. 9a).

The usefulness of thermal inertia information is in determining boundaries between materials of similar reflective properties but different thermal properties. An obvious example of these images is the differentiation between the basalt flows and the cinders. Both appear very dark in the albedo image (and in all ten reflective channels) but the cinders have much lower thermal inertia than the basalt. This can be seen in Fig. 9b and c where Pisgah crater, a cinder cone, clearly has a much lower thermal inertia than the surrounding lava flows.

Another application of thermal inertia which is illustrated by these images is the sensing of one material through a thin layer of another. There

are several areas on Lavic Lake where there is a few cm thick outwash of basalt overlying the playa material. In the albedo image (Fig. 4) these appear to be basalt. The thermal inertia of these areas, however, is either intermediate between the values for basalt and playa material, or is within the range of values for the playa. This effect can also be seen on the lower end of the Pisgah lava flow (Fig. 8). The areas of high thermal inertia (bright) correspond to solid basalt while the areas of lower thermal inertia have mixed sand and basalt debris, with a very dark desert pavement of basalt fragments on the surface.

Thus, using the image processing capabilities of JPL/IPL we have been able to produce and interpret images of a physical property, thermal inertia, from a diverse set of remotely sensed data. We will continue developing this interpretive technique for geologic applications.

REFERENCES

1. Wesselink, A. F., "Heat Conductivity and Nature of the Lunar Surface Material," Bull. Astron. Institutes Netherlands, X, 351-363, 1948.

2. Jaeger, J. C., "Conduction of Heat in a Solid with Periodic Boundary Conditions, With an Application to the Surface Temperature of the Moon," Proc. Cambridge Phil. Soc., 49 (2), 355-359, 1953.

3. Watson, K., "Geologic Applications of Thermal Infrared Images." Proc. IEEE, 63, 128-137, 1975.

4. Kahle, Anne B., "A Simple Model of the Earth's Surface for Geologic Mapping by Remote Sensing, " submitted for publication, 1976.

5. Joseph, J. H., Calculation of Radiative Heating in Numerical General Circulation Models, Numerical simulation of Weather and Climate Tech. Rpt. #1, Dept. of Meteor., Univ. of Calif., Los Angeles, 1966.

6. Joseph, J. H., "On the Calculation of Solar Radiation Fluxes in the Troposphere," Solar Energy, 13, 251-261, 1971.

7. Arakawa, A., Katayama, A., and Mintz, Y., "Numerical Simulation of the General Circulation of the Atmosphere," Proc. WMO/IUGG Symp. on Numerical Weather Prediction, Tokyo, 1968.

8. Kondratyev, K. Ya., Radiation in the Atmosphere, Academic Press, New York, 1969.

9. Burke, C. J., "Transformation of Polar Continental Air to Polar Maritime Air, J. of Meteor., 2, 94-112, 1945.

10. Sasamori, T., "A Numerical Study of Atmospheric and Soil Boundary Layers," J. Atmos. Sci., 27, 1122-1137, 1970.

11. Vehrencamp, J. E., "Experimental Investigation of Heat Transfer at an Air-Earth Interface," Trans. Amer. Geophys. Union, 34, 22-30, 1953.

12. Gillespie, Alan R., and Kahle, Anne B., "The Construction and Interpretation of a Digital Thermal Inertia Image," submitted for publication, 1976.
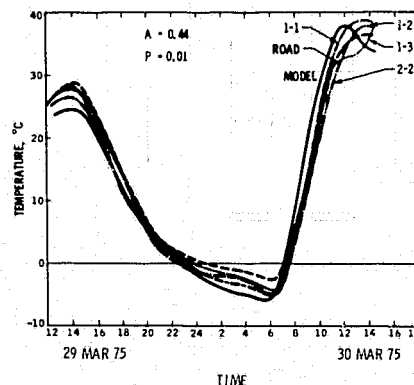
Fig. 1 Surface temperature from model and measurements
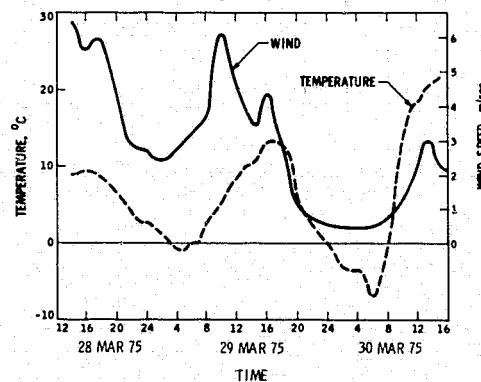
Fig. 2 Measured wind speed and air temperature
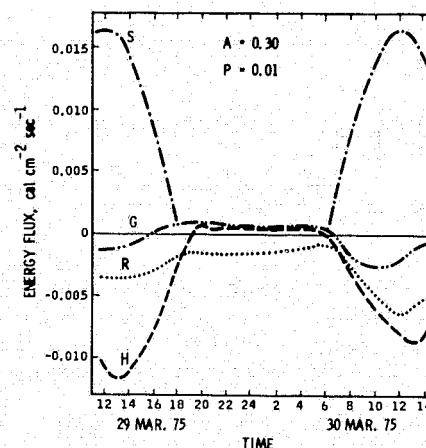
Fig. 3 Heating terms predicted by model
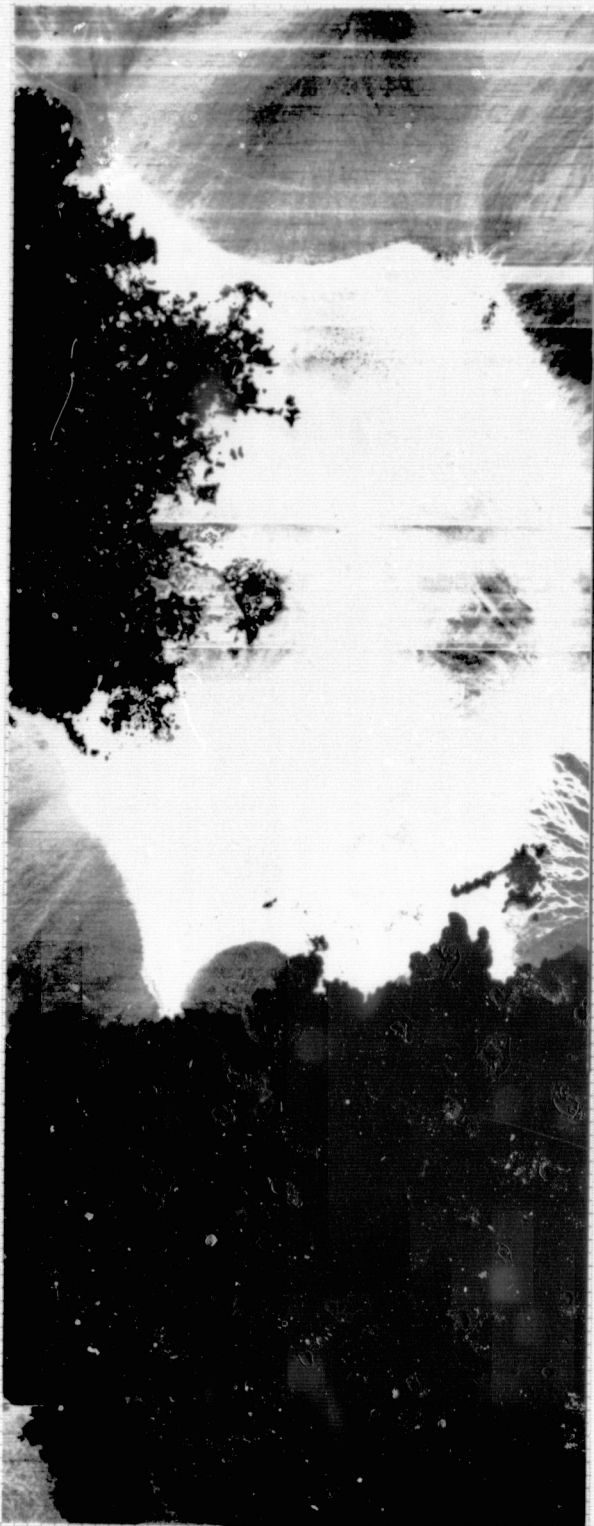
Fig. 4 Albedo
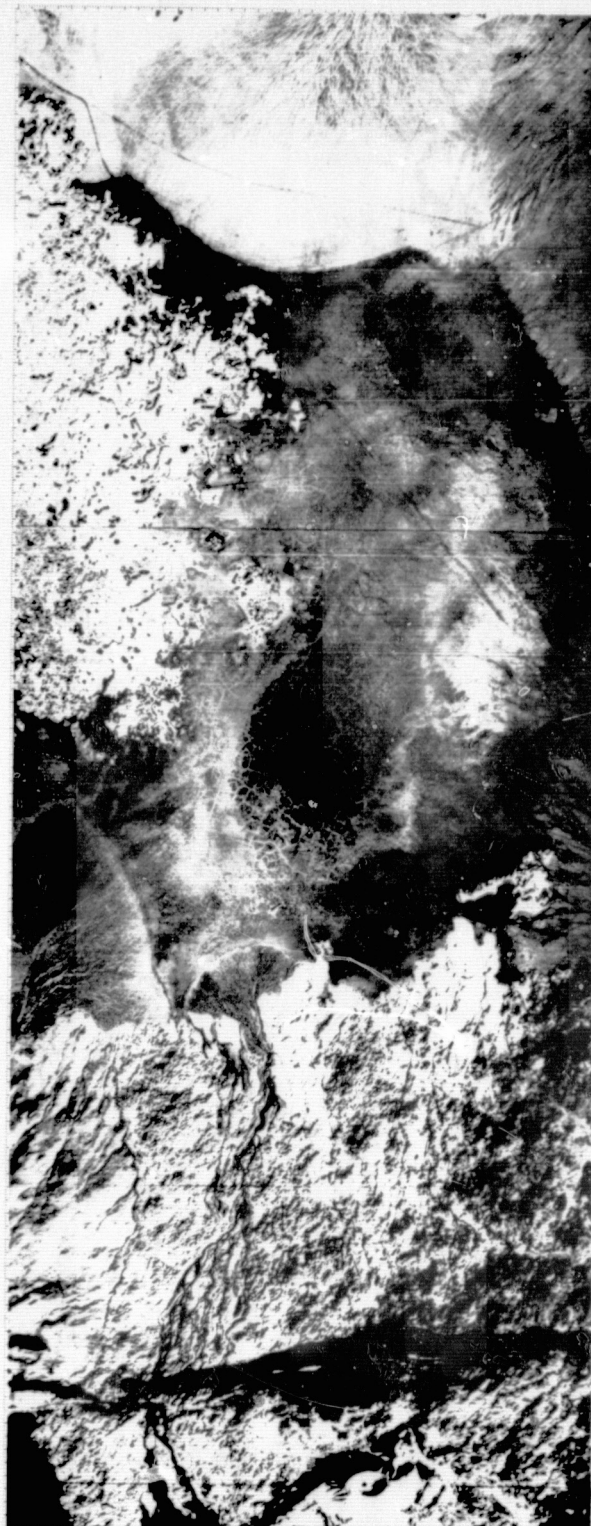
Fig. 5 Day Temperature
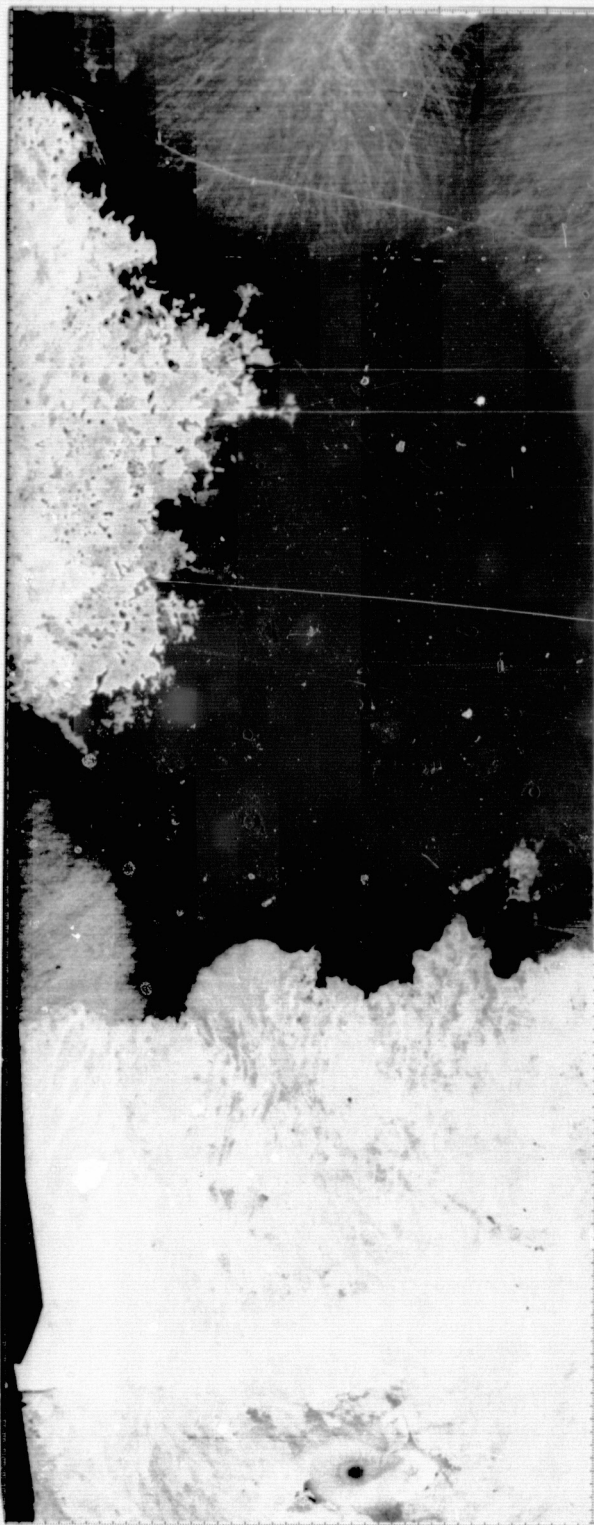
Fig. 6 Night Temperature

Fig. 7 Day Minus Night Temperature

Fig. 8 Thermal Inertia

11-7
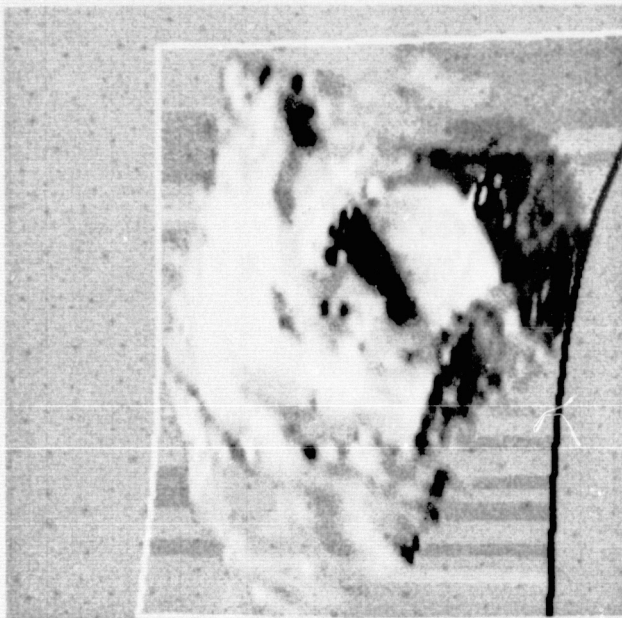
Fig. 9a) Computed insolation at Pisgah Crater

Fig. 9b) Thermal inertia without topographic correction
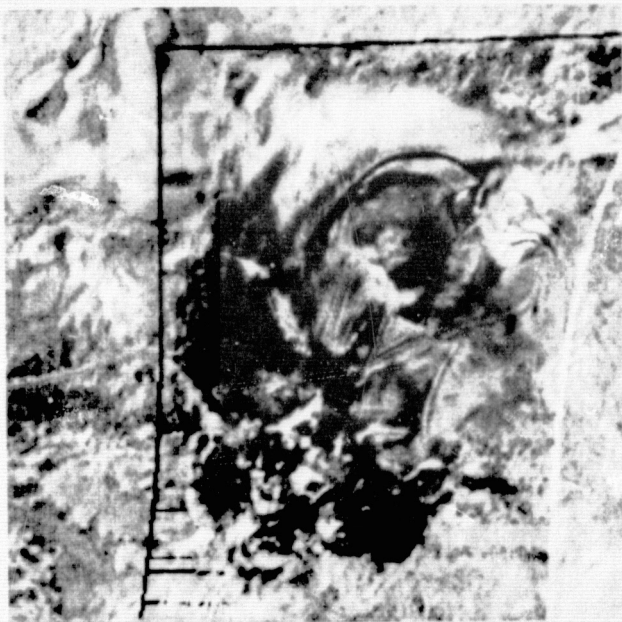
Fig. 9c) Thermal inertia with topographic correction
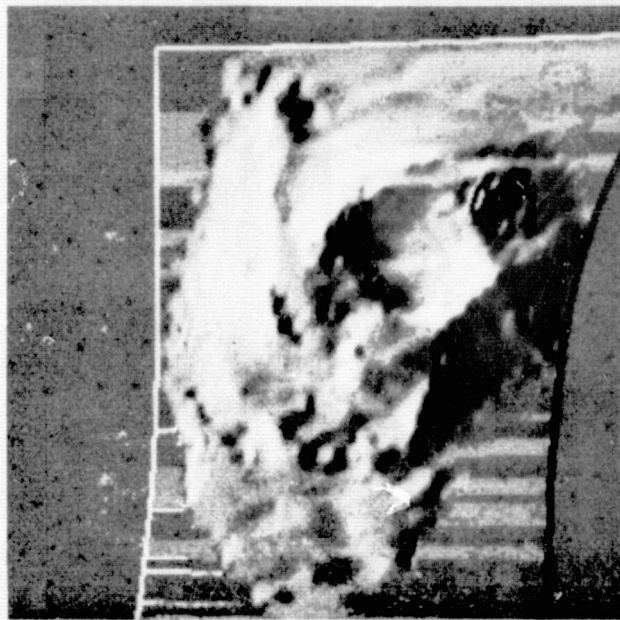
Fig. 9d) Change in calculated thermal inertia due to topographic correction

Peter R. Paluzzi

Jet Propulsion Laboratory
Pasadena, California

ABSTRACT

The application of computer image processing to ocean bottom photography and side-looking sonar marine resource imagery is now under study at the Jet Propulsion Laboratory. This report described the acquisition, processing, and interpretation of side-looking sonar records from San Francisco Bay for estuarine bedform recognition. It is shown that the processing and analysis of side-looking sonar imagery is possible using techniques developed for terrestrial and planetary remote sensing and that signatures in the processed records can be used to infer bottom conditions.

BACKGROUND AND DETERMINATIONS

Research sponsored by the National Aeronautics and Space Administration is now underway at the Jet Propulsion Laboratory (JPL) to apply data handling and analytical techniques used in planetary exploration to marine resource or seafloor imagery. Bottom photography and side-looking sonar are used to visually portray the seafloor, and application of JPL developed computer image processing to both forms of imagery has been described by Paluzzi and others (Ref. 1). This paper describes the methods used to collect, process, and interpret side-looking sonar records.

Side-looking sonar was developed around 1950 for military surveillance. The geologic use of this type of imagery was first reported by Chesterman and others (Ref. 2); since then, usage has increased steadily (Refs. 3-5). Briefly, side-looking sonar projects vertically downward and outward, normal to the path of the survey vessel. For each generated acoustic pulse, echos from the adjacent topography are registered (Fig. 1); soundings made along the ships track form an image of the seafloor. Belderson and others (Ref. 6) have described this technique, and Cole (Ref. 7) explains intrinsic problems and peculiarities of the side-looking method.

Side-looking sonar is key in the study of large seafloor areas since the overlying water does not permit long distance viewing or radar probing (Refs. 6, 8). Near bottom geology, geomorphology, mineral deposits, and geologic hazards over distances greater than 30 meters or so are best acoustically observed. Standard image display techniques provide a display of spatial relationships, surface features, and area texture with more detail than depth profiles and individual sample stations. The capacity to fill the space between survey tracklines is the most important property of seafloor imaging.

A strong resemblance exists between side-looking sonar "imagery" and scanned, remotely sensed images. The illumination geometry is essentially the same for side-looking radar and sonar, except for the different transmitting media and propagation speeds (Ref. 6). Since the sonar builds an image one line at a time along the ship's track a strip picture results, similar to that attained from airborne scanners. The equivalence between marine resource imagery and aerial remotely sensed imagery provides the opportunity to apply existing analytical methods to seafloor survey imagery.

It was determined by this study that the processing and analysis of side-looking sonar imagery is possible using techniques developed for terrestrial and planetary remote sensing. Side-looking sonar records processed by these techniques can be used to infer bottom conditions from image signatures.

IMPORTANCE TO USERS

Use of pre-existing image processing facilities, software and analytical techniques provide economy in the development of side-looking sonar as a survey tool similar to the economic benefits realized in the application of planetary image processing techniques to earth resources remote sensing (Ref. 9). Direct mapping of seafloor morphology, structures and deposits from sonar images means that surveys can be executed quickly, regionally and more completely. The importance of side-looking sonar as a data-source rests in its display of large seafloor areas.

In this study, the computer image processing of side-looking sonar records was successful because acceptable images were produced for interpretation. In the processed imagery, waveforms and other currents features are visible, as well as bright patchy return in some areas; both types of signatures are associated with hydrologic and sedimentary conditions in the survey area. Since a large 200m swath of seafloor is viewed instead of a standard depth profile, the amount of morphologic and textural information available to the interpreter is increased significantly for intertrack areas.

IMAGE ACQUISITION AND ANALYSIS

A side-looking sonar survey of a sand wave field was done on July 14 and 15, 1976, aboard the U.S. Geological Survey vessel the R/V Polaris, in San Francisco Bay. This survey covered a track 1200 meters long, oriented along 268°W and located 1300 meters north of the San Francisco waterfront (see Fig. 2). Selection of a single trackline was done to improve the changes for overlap in the records; however, problems with navigation and harbor traffic did not allow this with the 100 meter range sonar. An Edo-Western Model 606 side-looking sonar unit was used to perform the survey and display the records in real time; both port and starboard sonar signals were also recorded on an Ampex PR-2200 f.m. analog instrumentation tape recorder. Time-zero pulses and voice annotation were directly recorded on separate tracks. The analog tape recordings were digitized to 12 bits (4095 signal levels) upon returning to the shore and transferred to 9-track, computer compatible magnetic tape.

These computer tapes were used in the process-ing of the sonar images. All processing was per-formed at the Image Processing Laboratory (IPL) of JPL using the Video Image Communication and Retrieval (VICAR) image processing system on an IBM 360/65 computer; image playback was done on an Optronics International film reader. The computer handling of the side-looking sonar was performed in two stages, called first and second order process-ing. First order processing included conversion of the image data into VICAR compatible format and provided initial displays from the film recorder. The 12-bit digital data was reduced to 8 bits by dividing by 16, since the film recorder accepts only 256 intensity levels (8 bits). After this, a subset was selected for further analysis and inter-pretation. The sonar images suffered from a gro、、, cross-track, illumination fall-off or shading due to seawater signal loss and spherical spreading (Ref. 10); most side-scan recorders compensate for this with time varied gain (TVG) circuitry. The purpose of the second order processing was to suppress the shading effect with a different method: digital high-pass filtering. Planetary imagery from the Mariner and Viking missions is often high-pass filtered to enhance total scene detail and suppress low frequency (slowly varying) intensity variations (Refs. 11-13); this technique has been demonstrated successfully on seafloor photography and other side-looking sonar (Ref. 1). The high-pass filtered image represents the remain-der after local averages are subtracted from each picture element (pixel). The resultant high pass filter image is contrast enhanced to maximize the information displayed in the final products. The second order processing outputs were also scaled to 8 bits for playback.

Figure 3 shows a comparison between a side-looking sonar, underway chart record (Fig. 3a); film playback from the first order processing (Fig. 3b); and the result of high-pass filtering (Fig. 3c). Both forms of film playback do exhibit more detail than the original chart, yet the high-pass filtered version provides a more uniform presentation.

A complete side-looking sonar processing sequence appears in Fig. 4; note that for this report no geometric correction for slant range, ship speed, and heading was made due to time limi-tations and lack of adequate navigation data. Paluzzi and others (Ref. 1) have demonstrated that digital rectification is a useful and important aspect of computer image processing for side-looking sonar.

Figure 5 shows a larger stretch of side-looking sonar record. The vessel followed a course of 268° to true north and began a turn near a sunken wreck. There are two important bottom features that appear in this image: repetitive waveforms and patchy areas of stronger returns.

The waveforms or ripples are prominent pat-terns that cross the ship's track (Fig. 5); heights of the large waveforms vary from 1 to 5m with scene average of 3m; wavelengths vary from 28 to 88m with a mean of 60m. A smaller irregular set of ripples is superimposed on the larger waves. The large waveforms are asymmetric with broad gentle slopes facing into the path of the ship and steep sides facing down track. The tops of the waveforms are rounded with sharp slope breaks at

crests; troughs between the waves are flat. Both crests and troughs are straight, continuous, and sub-parallel with no bifurcations within the field of view. Alignment of the large waveforms is oblique to the ship's track with a measured inter-section of 12° to port or 10° to true north. The large waveforms were ubiquitous throughout the survey track; however, degraded waveforms were observed near point A. First returns from the fish nadir between points A to B clearly show the wave profile (see Figs. 1 and 5). The smaller waveforms are 3 to 4m in length with all under 0.5 meter in height.

A good view of the smaller waves appears in Fig. 5 at the point of turning. Here the course change was a turn to starboard and a new heading of 88° to true north. The crest of a large wave-form was in view during the turn, and it is dis-played deformed. Still visible are the broadly sloping side and the shadow caused by the steep break in slope. Forms on the broad, large, wave slopes are small waves. Since the illumination geometry of the side-looking technique causes shadows to fall away from the record's center, the small waves cast sharp, thin, dark shadows on the port side and diffuse wide, lighter shadows on the starboard side. This indicates asymmetry oppo-site of that presented by the larger waves. Other areas in the large wave field below have some superimposed smaller waves that cast shadows in a manner similar to the larger forms. Smaller wave orientations are subparallel to the large waves throughout these records and are absent from the large wave troughs.

Patchy areas of bright return are visible scattered throughout this portion of sonar record. Sizes are not consistent except in the large wave troughs where widths are in step with the wave-lengths. Other highly reflective areas are more irregular. For example, the two wispy patches in the lower left of Fig. 5 and the region of lighter gry at the upper right portion image. Very bright spots occur at the top left, to the left of the wreck, and far left of center. Except for the troughs, the borders of all bright patches are irregular. In some cases the edges are sharp; others are diffuse. None of the bright patches are completely homogeneous and in some areas, dark spots appear within their boundaries. Typical sizes for the bright patches vary as much as their shape. The smaller ones can be 2 to 5 meters across, while the oblong and larger patches are on the order of a hundred meters.

Constraints on survey time and navigation problems did not allow sampling and current measure-ments. Accurate navigation would have been needed if sampling was attempted on the feature within the area. In lieu of this, observations and interpre-tations from others (Ref. 14-17) were used as background.

Circulation in estuaries is predominantly by tidal currents (Ref. 18). In San Francisco Bay, the tidal circulation is modified by a jet effect that causes ebb over flood tidal current predomi-nance at the Golden Gate and greater flood currents near the survey area (Rubin, personal communication). U.S. Coast and Geodetic Survey current measurements at the Golden Gate and in the bay (stations 37 and 59 in Fig. 2) show this effect (see Table I).

The types of bedforms found in an estuarine, seafloor environment have been shown by flume studies (Ref. 19) and surveys (Ref. 14) to be controlled by current, grain size, and depth. Current ripples are a particular bedform that arise on non-cohesive bed surfaces due to a unidirectional current or fluid flow (Ref. 15). They can be either small or large depending on current velocities for a given grain size (Refs. 15, 20). Small current ripples are usually less than 30 to 60cm in length with wave height between 0.3 to 6 cm; the grain size of bed must be less than 0.5mm (Ref. 15). Large or mega-current ripples are mostly longer than 60cm (Ref. 20) with a maximum length of 30m; ripples longer than 30m are giant ripples (Ref. 15). Since bottom waveforms are created by a unidirectional flow and a tidal environment supports directionally varying currents, their direction and migration will depend on a net flow (Ref. 15). Stride points out that a 0.1 know difference between peak ebb and flood current velocities is sufficient to initiate migration of sandwaves. Table I shows that flood tidal velocities are much greater than the ebb currents inside the bay; this coupled with 0.1 to 2mm sediment grain sizes (Rubin, personal communication) indicates that if waveforms exist in the survey area, they would be sandwaves that migrate into the bay.

Urick (Ref. 10) has identified the physical parameters affecting acoustic reverberation and scattering from seafloor materials. Density, porosity and surface roughness are the most important factors controlling echo reflection from bottom sediment — depending on the wavelength. Dense materials will have a high acoustic impedance. The effect is the same for tightly packed sediments with little porespace. Scattering from the sediment water interface is proportional to the roughness and grain size. Coarse particles and disturbed beds will strongly scatter sonar echos with no specular reflection. Fine silts and muds have the opposite effect.

Correlations are possible when the side-looking sonar signatures were compared with the tide, current grain size and other data. Rubin and McCulloch (Ref. 14) have done this for much of San Francisco Bay using unprocessed chart records. The results have produced a model for estuarine bedform production. The interpretation presented here demonstrates that similar work can be done with computer processed imagery.

Waveforms that appear in the side-looking sonar records have shapes, sizes, and orientations consistent with megacurrent sandwaved produced by tidal currents. The predominance of tidal flood currents have produced megaripples in the uncon-solidated sand in the survey area that face into the bay (Ref. 14 and personal communication). Sets of smaller waveforms are due to slower cur-rents and are more responsive a directional change during a tide cycle (Ref. 15). Areas of smaller waves that are asymmetric opposite to the large waves are remnants of a previous ebb current.

Since no samples of the bright patches were taken, the exact cause of the strong returns is unknown. However, in some instances the size, shape, and location can be used to infer likely causes. The long linear patches with sharp border found in the wave troughs are probably scour lag surfaces that occur when the finer sediments

accumulate in the sand waves. Irregular patches that are randomly distributed are due to either isolated areas of coarse particles or seafloor dis-turbed by bioturbation. At times, artifacts like scrap metal and other hard objects will cause a strong echo.

FURTHER WORK

At present, the state of the art of seafloor imaging is still primitive relative to other remote sensing activities, particularly terrestrial and planetary applications. Signal acquisition is done by analog means, integration of ancillary data (speed, navigation, orientation, etc.) is at best awkward, a preponderance of post-acquisition image corrections are needed, and the display media (electro-sensitive paper) is appalling by current remote sensing standards.

Any effect to address these problems will result in better and more useful images. A shift in processing emphasis from image correction to information extraction will allow analysts and scientists to pursue large area survey, pattern analysis, grain size and texture measurement from imagery, and automated deposit mapping.

In terms of future San Francisco Bay surveys, improved sonar imagery would permit high resolution coverage of larger areas to extend observation of sedimentary processes. Rapid survey with good spatial registration will permit change detection for mean numbers of tidal and seasonal cycles.

The application of existing remote sensing processing and analysis to seafloor exploration will require more study as we enter an age of increased resource and environmental awareness.

ACKNOWLEDGEMENTS

REFERENCES

1. Paluzzi, P. R., Normack, W. R., Hess, G. R., Hess, H. D., Cruickshank, M. J., "Computer Image Processing in Marine Resource Exploration," Proceedings for Joint IEEE-MTS Oceans '76 Conference, Washington, DC, September 13-15, 1976.

2. Chesterman, W. D., Clynick, P. R., Stride, A.H., "An Acoustic Aid to Sea-bed Survey," Acustica, Vol. 8, pp. 280-290, 1958.

3. Donovan, D. T., and Stride, A. H., "An Acoustic Survey of the Sea Floor South of Dorset and Its Geological Interpretation," Philosophic Trans-actions of the Royal Society of Britain,

Vol. 244, No. 712, pp. 299-330, 1961.

4. Sanders, J. E., Emery, K. O., and Uchupi, E., "Microtopography of Five Small Areas of the Continental Shelf by Side-Scanning Sonar," Geological Society of America Bulletin, Vol. 80, pp. 561-572, 1969.

5. Mudie, J. D., Normark, W. R., and Gray, E. J., Jr., "Direct Mapping of the Seafloor Using Side-Scanning Sonar and Transponder Navigation," Geology Society of America Bulletin, Vol. 81, pp. 1547-1554, 1970.

6. Belderson, R. H., Kenyon, N. H., Stride, A. H., and Stubbs, A. R., Sonographs of the Sea Floor, A Picture Atlas, Elsevier, New York, 1972, 185 p.

7. Cole, F. W., A Familiarization with Lateral or Side-Scanning Sonars, Columbia University, Hudson Laboratories, Dobbs Ferry, NY, 1968.

8. Menard, H. W., Marine Geology of the Pacific, McGraw-Hill, New York, 1964, 271 p.

9. Goetz, A.F.H., Billingsley, F. C., Gillespie, A. R., Abrams, M. J., Squires, R. L., Shoemaker, E. M., Luchitta, I., Elston, D. P., "Application of ERTS Images and Image Processing of Regional Geologic Proglems and Geologic Mapping in Northern Arizona," Technical Report 32-1597, Jet Propulsion Laboratory, Pasadena, CA, 1975.

10. Urick, R. J., Principles of Underwater Sound, McGraw-Hill, New York, 1975, 384 p.

11. Green, W. B., Jepsen, P. L., Kreznar, J. E., Ruiz, R. M., Schwartz, A. A., and Seidman, J. B., "Removal of Instrument Signature from Mariner 9 Television Images of Mars," J. Appl. Optics, Vol. 14, pp. 105-114, 1975.

12. Seidman, J. B., "Some Practical Applications of Digital Filtering in Image Processing," Proceedings of the Symposium of Computer Image Processing and Recognition, Vol. 2, Dept. of Electrical Engineering, University of Missouri, Columbia, MO, 1972.

13. Soha, J. M., Lynn, D. J., Lorre, J. J., Mosher, J. A., Thayer, N. N., Elliott, D. A., Benton, W. D., and Dewar, R. E., "IPL Processing of the Mariner 10 Images of Mercury," J. Geophys. Res. Vol. 80, pp. 2394-2414, 1975.

14. Rubin, D. M., McCulloch, D. S., "Bedform Dynamics in San Francisco Bay, California," Geological Abstracts, in press.

15. Reineck, H. E., Singh, I. B., Depositional Sedimentary Environments, Springer-Verlag, New York, NY, 1973, 439 p.

16. Werner, F., and Newton, R. S., "The Pattern of Large Scale Bedforms in the Langeland Belt (Baltic Sea)," Marine Geology, Vol. 19, pp. 29-59, 1975.

17. Stride, A. H., "Current Swept Seafloors Near the Southern Half of Great Britain," Quarterly Journal of the Geological Society of London, Vol. 119, pp. 175-199, 1963.

18. Gorsline, D. S., "Contrasts in Coastal Bay Sediments on the Gulf and Pacific Coasts," in Estuaries, Pub. No. 83, American Association for the Advancement of Science, Washington, DC, 1967, 757 p.

19. Simons, D. B., and Richardson, E. V., "Resistance to Flow in Alluvial Channels," American Society of Civil Engineers, Transactions, Vol. 127, pp. 927-953, 1962.

20. Allen, J. R. L., Current Ripples. Their Relation to Patterns of Water and Sediment Motion, North-Holland Publishing Company, Amsterdam, 1968, 433 p.
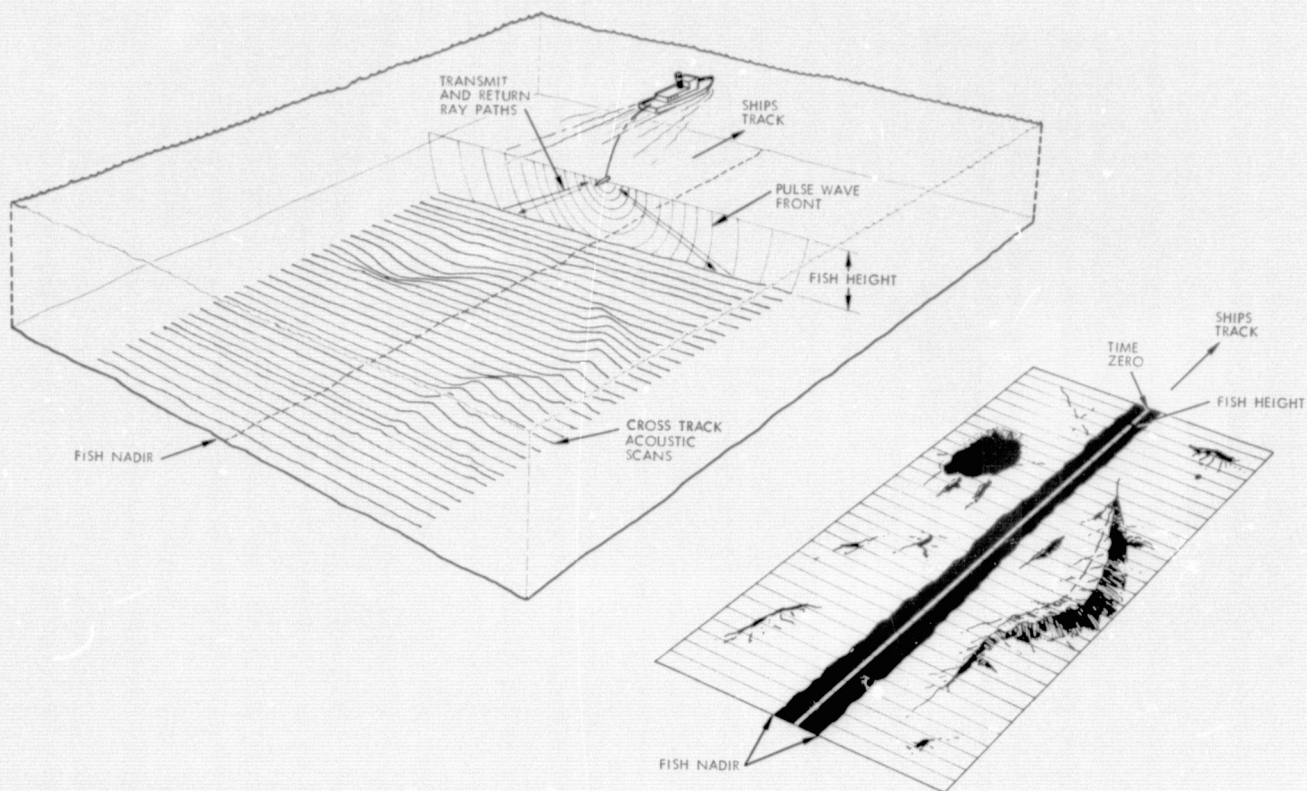
Figure 1. Schematic view of side-looking sonar, imaging technique. A sonar transceiver is towed by a survey vessel and cross-track acoustic scans are made (a). These scans are recorded aboard the vessel and displayed as a chart (b).
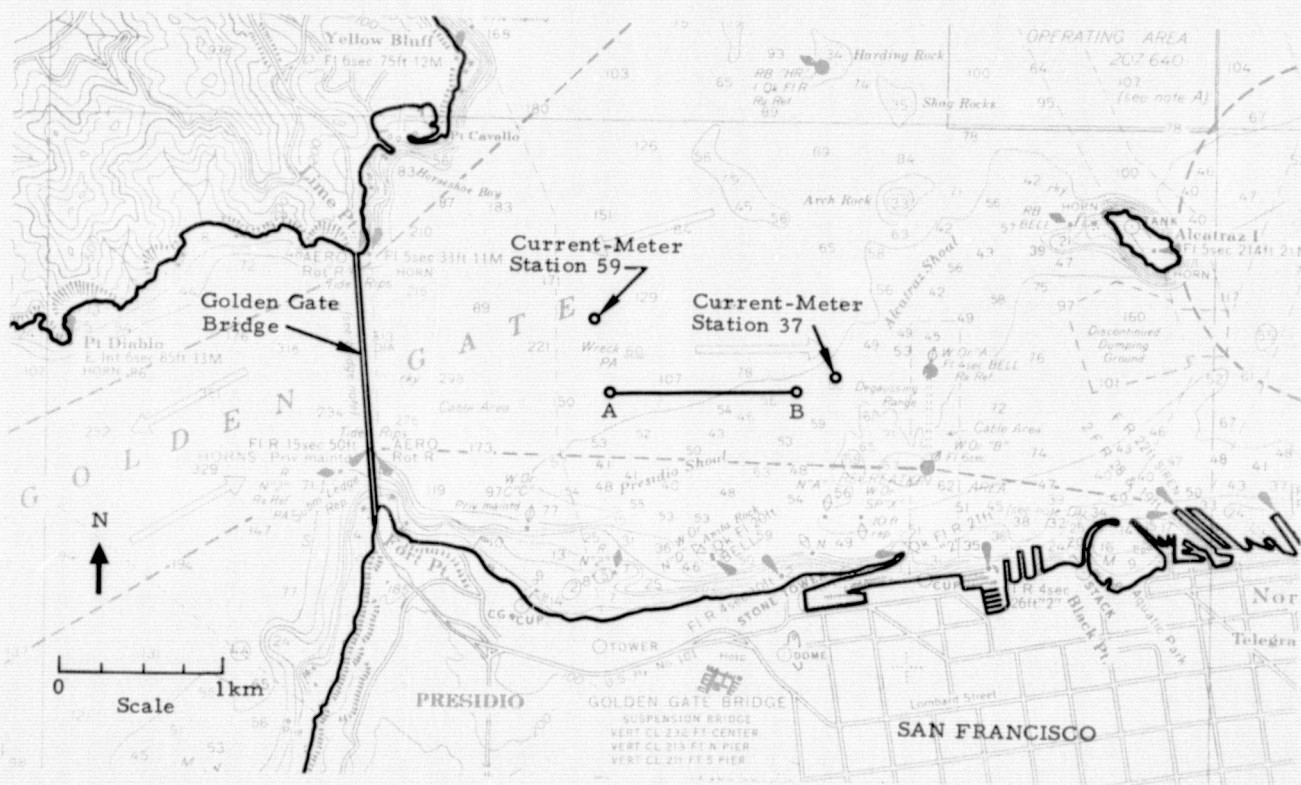
Figure 2. Location of side scan survey. Six coinciding tracks were run between points A and B. After NOAA/National Ocean Survey, Nautical chart number 18649.
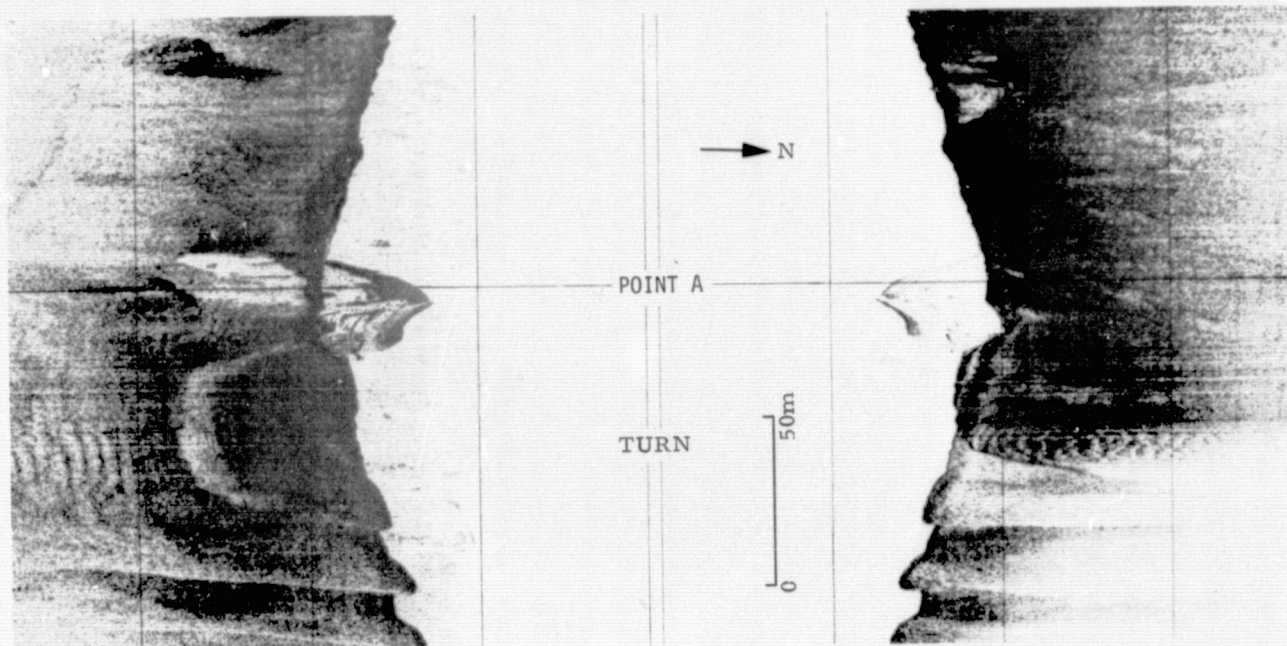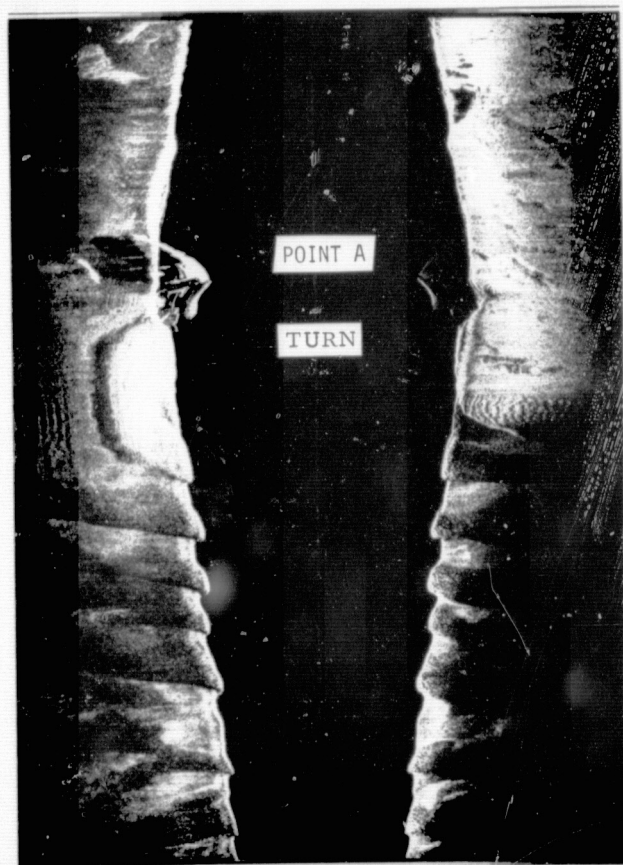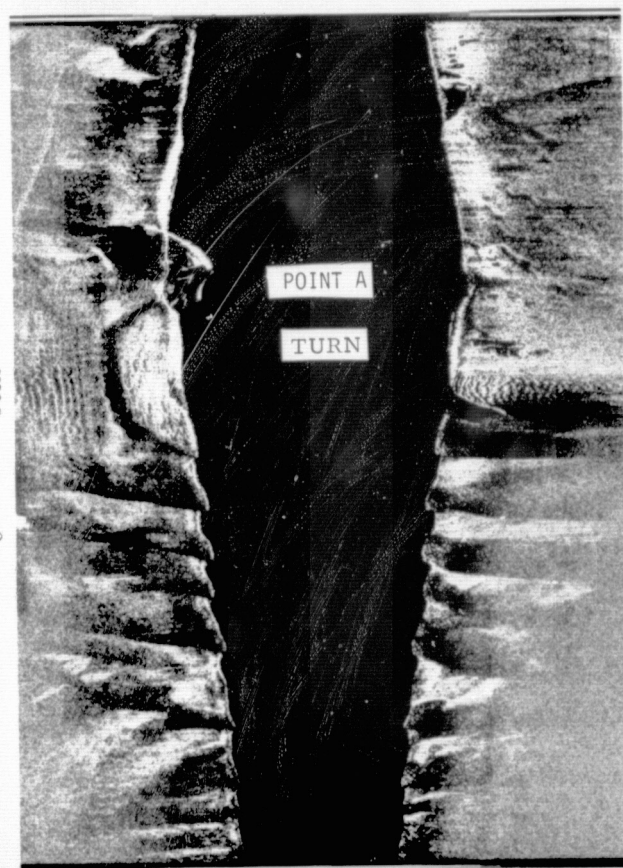
Figure 3. A portion of side-looking sonar record going from point A to B
as: the original chart record (a); the first order playback (b);
and the processed final playback (c).

```
┌──────────────┐              ┌──────────────┐
│ RECORD       │              │ COMPUTE SLANT│
│ ANALOG       │              │ RANGE        │
│ DATA         │              │ CORRECTION   │
└──────────────┘              └──────────────┘
        │                            │
┌──────────────┐              ┌──────────────┐
│ DIGITIZE AND │              │ CORRECT      │
│ MAKE COM-    │              │ IMAGE FOR    │
│ PUTER TAPE   │              │ SLANT RANGE  │
└──────────────┘              └──────────────┘
        │                            │
┌──────────────┐              ┌──────────────┐
│ LOG DATA     │              │ CORRECT FOR  │
│              │              │ SHIP'S SPEED │
└──────────────┘              └──────────────┘
        │                            │
┌──────────────┐                    ◇
│ FILTER FOR   │              IS NAVIGATIONAL    YES   ┌──────────────┐
│ VERTICAL     │              DATA AVAIL- ABLE ?  ───→ │ CORRECT FOR  │
│ HIGH PASS    │                    ◇                  │ HEADING      │
└──────────────┘                    │ NO              └──────────────┘
        │                    ┌──────────────┐                │
┌──────────────┐             │ CONVERT FROM │                │
│ FILTER FOR   │             │ VIDEO TO FILM│←───────────────┘
│ HORIZONTAL   │             └──────────────┘
│ HIGH PASS    │                    │
└──────────────┘             ┌──────────────┐
        │                    │ OBTAIN PRINTS│
┌──────────────┐             │ THROUGH      │
│ SCALE FOR    │             │ PHOTO LAB    │
│ 256 GRAY     │             └──────────────┘
│ LEVEL        │
└──────────────┘
        │
┌──────────────┐
│ REMOVE BAD   │
│ LINES AND    │
│ GLITCHES     │
└──────────────┘
        │
┌──────────────┐
│ DESPIKE TO   │
│ REMOVE       │
│ NOISE        │
└──────────────┘
```

Figure 4. Typical computer image processing sequence for side-looking sonar.
           NOTE: Correction for slant range, ship's speed, and heading were
                  not done for this study (see Ref. 1).
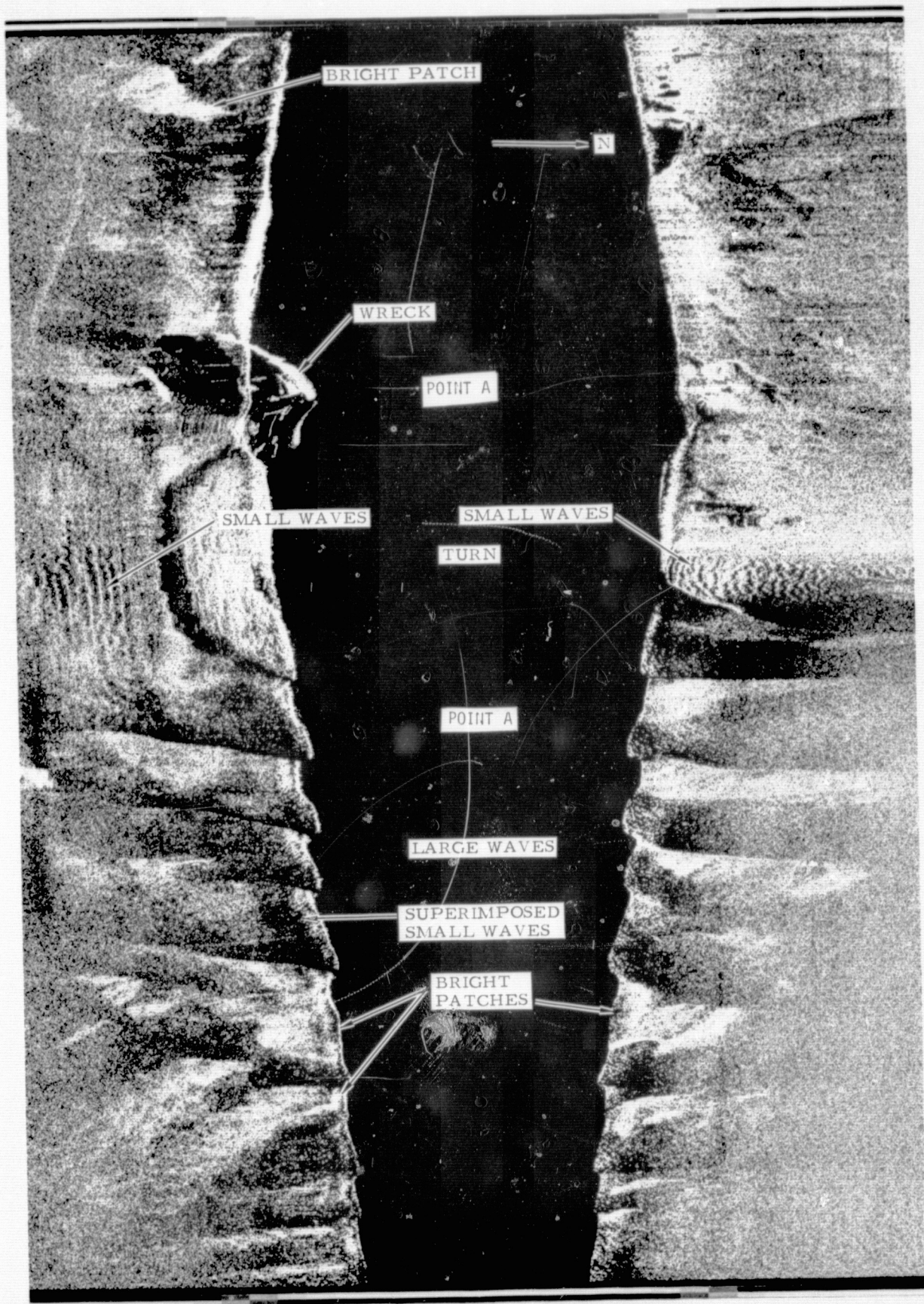
Figure 5. An expanded view of the image in Figure 3c.

Table I

U.S. Coast and Geodetic Survey
Current Velocity Data

Station*	Meter depths (ft.)	Mean flood velocity (knots)	Flood duration (hrs)	Maximum observed flood velocity (knots)	Mean ebb velocity (knots)	ebb duration (hrs)	Maximum observed ebb velocity (knots)
37	8	2.2	5.6	3.4	2.4	6.2	3.2
	27	2.2	6.7	3.3	2.1	6.1	3.0
	48	1.5	7.6	2.4	1.3	5.8	2.4
59	8	2.5	6.1	4.3	2.3	6.3	4.0
	85	2.5	7.0	4.3	1.8	5.5	2.9
	105	1.9	0.7	3.6	1.9	5.8	3.0
Golden Gate	8	3.7	--	4.7	4.0	-	5.9

*See Fig. 2

RECENT APPLICATIONS OF DIGITAL PROCESSING TO PLANETARY SCIENCE

Donald J. Lynn

Jet Propulsion Laboratory
Pasadena, California

ABSTRACT

During the past several years, digital images of all of the planets from Mercury to Uranus have been processed at JPL's Image Processing Laboratory (IPL). These include images from flyby missions, orbital missions, unmanned landers and ground-based observations. The primary objectives of all of this processing has been to enhance or extract information in support of a specific scientific investigation.

This paper presents selected examples from the Mariner 10 mission to Venus and Mercury, and the Viking mission to Mars. Mercator projections of far encounter images were made and mosaicked in support of the Venus Cloud Motion Studies; these projections required the development of a special "Far Encounter Algorithm" for computing the required transformation matrices and vectors. High quality, large scale photomosaics and computer mosaics were produced to help define the morphology of the surface of Mercury. Production of these mosaics required the development of new radiometric decalibration techniques, a "scene dependent" filter algorithm, and the refinement of computer mosaicking techniques. Precision stereo photogrammetric analyses of the surface of Mars are being performed on images from Viking Orbiter using an interactive capability developed specifically for that purpose. The topology of the local surface surrounding each of the Mars Viking Landers is being systematically mapped using a special hardware and software system developed for that purpose in conjunction with the Stanford Artificial Intelligence Laboratory. References are made to analyses performed on data from the outer planets.

INTRODUCTION

During the past several years, digital images of all of the planets from Mercury to Uranus have been processed at JPL's Image Processing Laboratory (IPL). These images have been obtained from flyby missions, orbital missions, unmanned landers, and earth-based observations. The primary objectives of all of this processing has been to enhance or extract information in support of a specific scientific investigation.

This paper presents a few examples of new software or new techniques developed for a particular scientific objective.

VENUS CLOUD MOTION STUDIES

In February 1974, the Mariner 10 television cameras imaged the planet Venus in the visible and near ultraviolet for a period of eight days. One of the scientific objectives of the imaging sequences taken during this period was to investigate the spatial and temporal characteristics of UV markings which had been observed from earth.

It was desired to provide two different types of analysis of cloud development as a function of time — one in which the behavior of specific features are followed around the planet, and a second in which the development of cloud features at a specific point are monitored as a function of time.

There were several factors which made direct comparison of images impossible for this purpose: (a) As the spacecraft receded from Venus (Fig. 1) the effective resolution in the images decreased from ~2 km to >125 km; (b) The contrast by which the features were distinguishable was considerably less than the global shading; (c) Even the a posteriori scan platform pointing information from the spacecraft was not good to more than about 100 pixels, and (d) Since only clouds were observable, there were no fixed tiepoints or "landmarks" with which to register images. Some technique had to be devised to permit identifying the latitude and longitude of each pixel in each of the images.

The technique devised was the "Far Encounter Algorithm: described by Denis Elliott (Ref. 1). Briefly, the algorithm takes advantage of (a) the fact that the ephemerides of the spacecraft and planet are known to significant accuracy; (b) the subspacecraft point is coincident with the center of the planet in the image; and (c) except for a full frontal view, the north angle of the planet can be determined from the orientation of the terminator. From this information, the appropriate matrices and vectors can be calculated and a latitude and longitude determined for each pixel.

For the Venus cloud motion studies, the information described above was input to the MAP2 program, and Mercator projections of each image were produced. In addition, the images were high pass filtered with a filter whose size was varied in order to maintain a constant relationship with the resolution of the image. These projected images were then cut up in small strips and mosaicked in such a way as to show cloud development at a specific point as a function of time.

Figure 2 is an example of one of the mosaics produced in this manner. With the aid of these mosaics, the lifetime and propagation characteristics of large scale and smaller scale markings was determined (Ref. 2).

MERCURY SURFACE MORPHOLOGY

A significant factor in the analysis of Mercury surface morphology from Mariner 10 data was the ability to construct high quality, large scale mosaics. Several developments in IPL contributed to this capability.

The radiometric decalibration techniques utilized in the Mariner 10 mission (Ref. 3) were similar to those used in the Mariner 9 mission (Ref. 4), with some small, but significant differences. One of these differences was a change in

the software which permitted images to be automatically scaled to a user specified decalibration constant (microwatts of solar spectrum energy per square centimeter per steradian per DN). Thus areas of equal brightness, if decalibrated with the same constant, should have the same DN values in the decalibrated images. This proved to be the case.

The initial proof of the validity of the decalibration procedure came when we were asked to produce the mosaic shown in Fig. 3. This is a low resolution photomosaic of images from the incoming leg of the Mercury I mission. It contains both UV and orange images. Compensation for differences in spectral reflectance was accomplished by merely determining and applying appropriate decalibration constants for each spectral filter. The mosaic shown was processed and assembled within 24 hours after receipt of the data. Having this and subsequent high quality global mosaics available at an early time in the mission was a significant benefit to the scientists in their geologic interpretation (Ref. 5).

Another significant technique developed to aid in the interpretation of Mariner 10 images was the use of a "scene dependent" filter. This was a spatial filter designed specifically for images taken in the vicinity of the terminator. The mathematical formulation for this filter is given more explicitly in Ref. 6, but a brief description is as follows. Near the terminator the illumination level goes to zero, thus, although the relative contrast in the scene stays about the same, the actual differences in data values goes to zero. Conventional high pass filtering or simple contrast enhancement techniques cannot compensate for this apparent degradation in contrast and, correspondingly, in discriminability. To overcome this, Jim Soha devised a filter algorithm in which the local contrast is amplified in inverse proportion to the local average scene brightness. This technique produces an image which simulates a uniformly illuminated scene.

The results of the application of this filter to one of the more outstanding geologic features on Mercury is shown in Fig. 4. As an added bonus, it turned out that this filter could be incorporated easily into one of the standard recursive high pass filtering algorithms used in IPL and has been utilized very effectively in several other applications.

The success achieved in producing high quality photomosaics led us to undertake the production of computer mosaics. This was made possible and practical by a number of factors which included: almost identical illumination conditions on all three Mercury encounters, high quality of the images produced by the imaging systems; the success of the decalibration procedures, the generation of accurate OM matrices and $\vec{OC}$ vectors (Ref. 1) by Merton Davies of the Rand Corporation, the development of the cartographic projection techniques described in Ref. 1, appropriate funding from the Mariner 10 project and the Planetoloty Program Office of NASA, and the capabilities and enthusiasm of Joel Mosher.

Initially, Joel produced the computer mosaic of the Caloris basin shown in Fig. 5. This mosaic was composed of fifteen images and incorporated the radiometric decalibration and scene dependence

filtering techniques described earlier and some new mosaicking algorithms developed by Joel.

That product proved to be so attractive and useful that Joel kept expanding his accomplishments until he produced products like that shown in Fig. 6. Figure 6 is a computer mosaic of 76 Mariner 10 images. This represents an input assemblage of over 44×10^6 bits. The output format is 7000 by 5000 pixels.

To date, Joel has produced computer mosaics of all the mapping quadrants of Mercury for which data is available. These mosaics are currently being used by the geologic mappers and will appear in a Mercury atlas which is scheduled to be completed early next year.

STEREOPHOTOGRAMMETRY FROM VIKING ORBITER

One of the scientific objectives of the Viking Orbiter mission is the analysis of surface topography based on stereo imagery. The stereo effect is obtained by twice imaging the same area on the surface as the spacecraft passes overhead. One image is recorded in a "forward looking" geometry, where the camera system is pointed at a location ahead of the spacecraft position, and the second image is recorded in a "backward looking" geometry, after the spacecraft has passed over the location of interest. Figure 7 shows a stereo pair recorded in this manner.

Determination of the relative displacement of a common point between the two digital images enables calculation of the three-dimensional coordinates of that point in some established coordinate system. A limited amount of analytic photogrammetric analysis was performed on a small number of images of Mars from Mariner 9 (Ref. 7) and on some of the Mariner 10 images of Mercury.

An interactive capability for performing precision digital stereo photogrammetry is being developed at IPL under the cognizance of Reuben Ruiz with support from the Viking Project. The software for this effort was developed by Jean Lorre, Gary Yagi, and Margie Power. The steps currently involved in producing a topographic map from a stereo image pair are as follows.

The analyst views the stereo pair on a single split image video display with a graphics overlay plane. The analyst identifies a feature in the left hand image using a trackball and cursor, then identifies the same feature in the right hand image using a second cursor. A cross correlation of the areas surrounding the cursored points is then performed to refine the analyst's positioning of the second cursor. From this correlation the location of the corresponding point in the right hand image is obtained with a precision of less than one picture element. In order to overcome problems with intensity variations caused by differences in illumination or viewing (e.g., the two images may have been taken at different times of the Martian day), the correlation is actually performed between the Fourier transforms of the image segments. This represents a significant improvement over methods previously employed.

The locations of many features in both the left and right images are determined in this manner, and the graphics plane is utilized to indicate the

points selected and the relative displacement vector between the locations of the same feature in both images. A typical displacement diagram is shown in Fig. 8.

The elevation of each cursored point is then computed, and an elevation image is created in which elevation variation in the region is converted to gray level intensities ranging from 0 to 255. Elevation "intensities" for the remaining picture elements are then computed by interpolation between the cursored points. The elevation "image" is then contoured, with the analyst selecting the contour intervals to correspond to the desired precision of topographic detail. The contour lines are then overlaid on one of the two images of the stereo pair. Figures 9 and 10 indicate 50m and 100m iso-elevation contours overlaid on the right image.

The number of points required for adequate mapping of a single stereo pair depends on the complexity of the terrain, the size of the images and the desired precision of the elevation contours. Typically for a Viking Orbiter image which is approximately 1200x1000, between 200 and 500 points must be cursored. This currently takes from 30 minutes to four hours of interaction to produce elevation data for a single pair.

MARTIAN SURFACE CHARACTERISTICS

One of IPL's primary responsibilities in support of the Viking mission has been that of providing stereo ranging support to the surface sample acquisition activities of the two Viking Landers. Each Lander has a pair of facsimile type cameras (Ref. 8) which provides stereo imaging of much of the surrounding area.

A special display and viewing station was designed by Sid Liebes of Stanford and built by the Stanford Artificial Intelligence Laboratory to permit stereoscopic viewing of the Lander stereo images.

The interactive program RANGER was developed by Arnie Schwartz of IPL in cooperation with Liebes to display the images, provide a constrained cursor in each image, and perform all photogrammetric computations required to determine the three-dimensional location of each cursored point in a specified coordinate system. The program has a variety of modes, but this discussion is limited to the ability to draw vertical profiles.

A typical scenario for use of this system is as follows. Appropriate segments from a stereo pair are displayed on the two monitors of the viewing station. The optics and displays of the station are adjusted to permit the photogrammetrist to fuse the two images. A cursor in each image is then moved with a trackball until it appears to coincide with a desired point in the images. The three-dimensional coordinates of that point are then computed from the image plane coordinates in each image. Corrections for camera geometric distortions are provided using camera-simulating software developed by Mike Wolf (Ref. 8). This software duplicates, as closely as possible, an actual ray trace through the camera systems and includes camera dependent variables to compensate for mechanical, optical, and manufacturing differences between cameras. The coordinates from each

validated cursored point are stored, and the point is displayed in the graphics plane of each image.

In the profile mode of RANGER, the cursor is mathematically constrained to movement in a specified plane; e.g., a vertical plane passing through the surface sampler arm axis. As the photogrammetrist moves the cursor along the apparent surface, the coordinates of each point are computed and stored. The resulting set of coordinate values, termed a range data set, may then be made available to programs which display the range data set as a contour map, or profile plot.

Figure 11 shows a typical Viking Lander stereo image segment with the profile traces superposed over the image. Figure 12 shows a typical plot of one of these profiles, produced using software developed by Rodger Philips. Profiles of this type have been made to characterize the sample site before and after sample collection. Full scale plots have been generated to permit accurate modeling of the Martian surface in the immediate vicinity of the Lander. Profiles to the horizon have been generated and used to characterize the general surrounding terrain.

SUMMARY

A few techniques developed at IPL in response to specific scientific objectives of NASA's Planetary Exploration Program have been presented. These are by no means inclusive or exhaustive. More detailed descriptions of developments for the Viking Project will be forthcoming. A description of some of the new techniques developed for analysis of earth-based planetary images may be found in Ref. 10.

ACKNOWLEDGEMENT

REFERENCES

1. Elliott, D. A., "Digital Cartographic Projection," in Proceedings of Caltech/JPL Conference on Image Processing Technology, Data Sources and Software for Commercial and Scientific Applications, Pasadena, CA, November 3-5, 1976.

2. Belton, M.J.S., et al., "Space-Time Relationships in the UV Markings on Venus," Journal of Atmospheric Sciences, to be published.

3. Soha, J. M., et al., "IPL Processing of the Mariner 10 Images of Mercury," Journal of Geophysical Research, Vol. 80, No. 17, June 1975.

4. Green, W. B., et al., "Removal of Instrument Signature from Mariner 9 Television Images of Mars," Applied Optics, 14, 105-114, 1975.

5. Murray, B. C., et al., "Mercury's Surface: Preliminary Description and Interpretation from Mariner 10 Pictures," Journal of Geophysical Research, 80, 1975.

6. Soha, J. M., "Scene Dependent Filtering for Variable Contrast Enhancement," Symposium on

Automatic Computation and Control, Milwaukee, WI, April 1975.

7. Blasius, K. R., "A Study of Martian Topography by Analytic Photogrammetry," _Journal of Geophysical Research_, _78_, 1973.

8. Wolf, M. R., "The Analysis and Removal of Geometric Distortion from Viking Lander Camera Images," presented at the Fall Convention of the American Society of Photogrammetry, October 1975.

9. Schwartz, A. A., "RANGER — An Interactive Stereo Photogrammetric Computer Program for the Viking Mars Lander Mission," presented at the Fall Convention of the American Society of Photogrammetry, October 1975.

10. Lorre, J. J., et al., "Recent Developments at JPL in the Application of Image Processing Techniques to Astronomical Images," _Proceedings of SPIE Asilomar Conference on Image Processing_, February 1976.
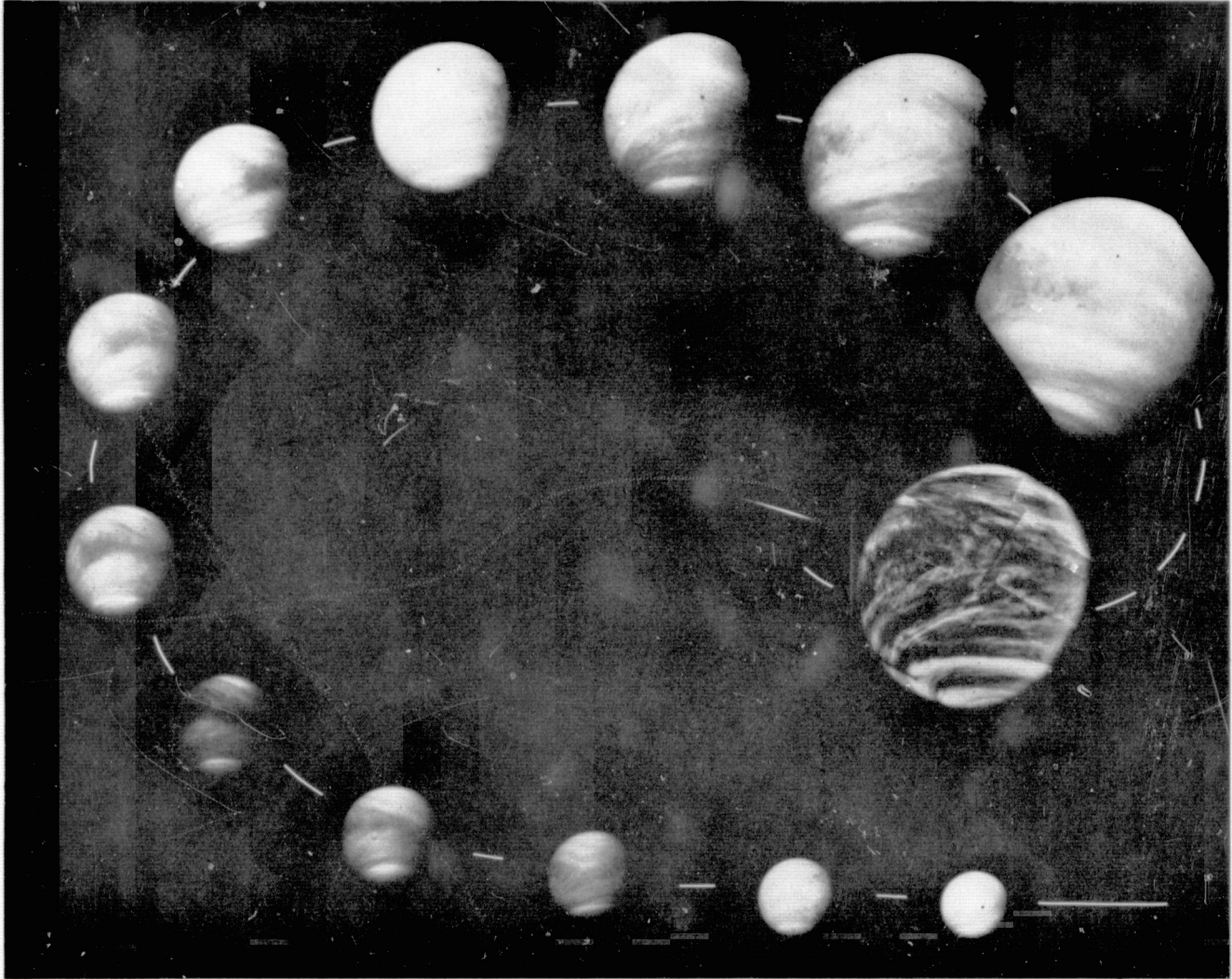
Figure 1. An assemblage of pictures of Venus taken in the UV from Mariner 10, illustrating the changes in resolution and scene content as a function of time. The largest is a photomosaic of high pass filtered images taken with a resolution of approximately 6 km. The next eight were taken at approximately 12-hour intervals with resolution degrading from 17 km to approximately 72 km. The last four were taken at 24-hour intervals with resolution degrading from 88 km to approximately 125 km.

(P-15552B)
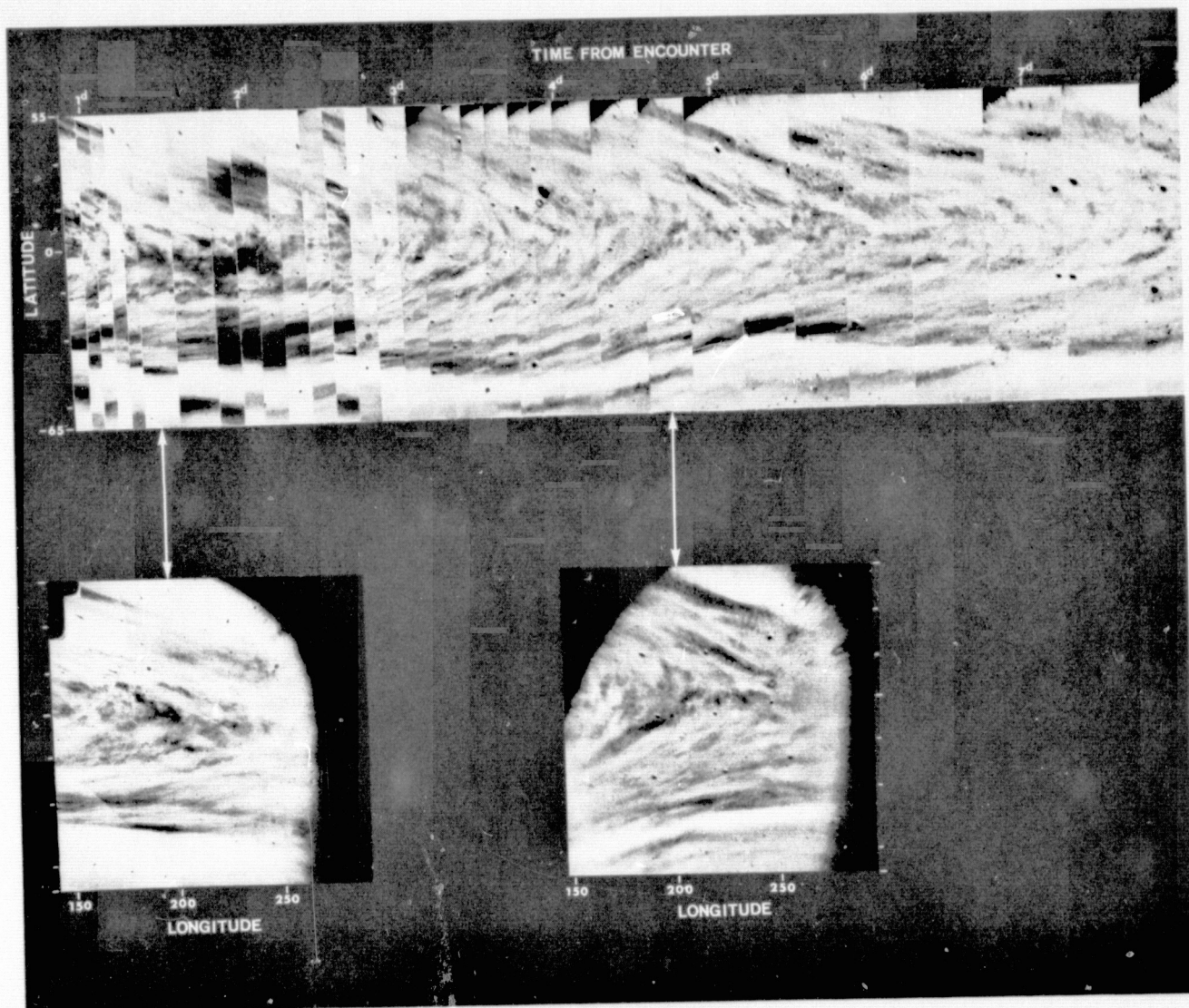
Figure 2. Upper — Development of UV markings as a function of time at the subpolar meridian of Venus. This is a photomosaic of strips of high pass filtered versions of images covering a period of 7 days.

Lower — Examples of two Mercator projections from which the narrow strips were taken (from Ref. 2).
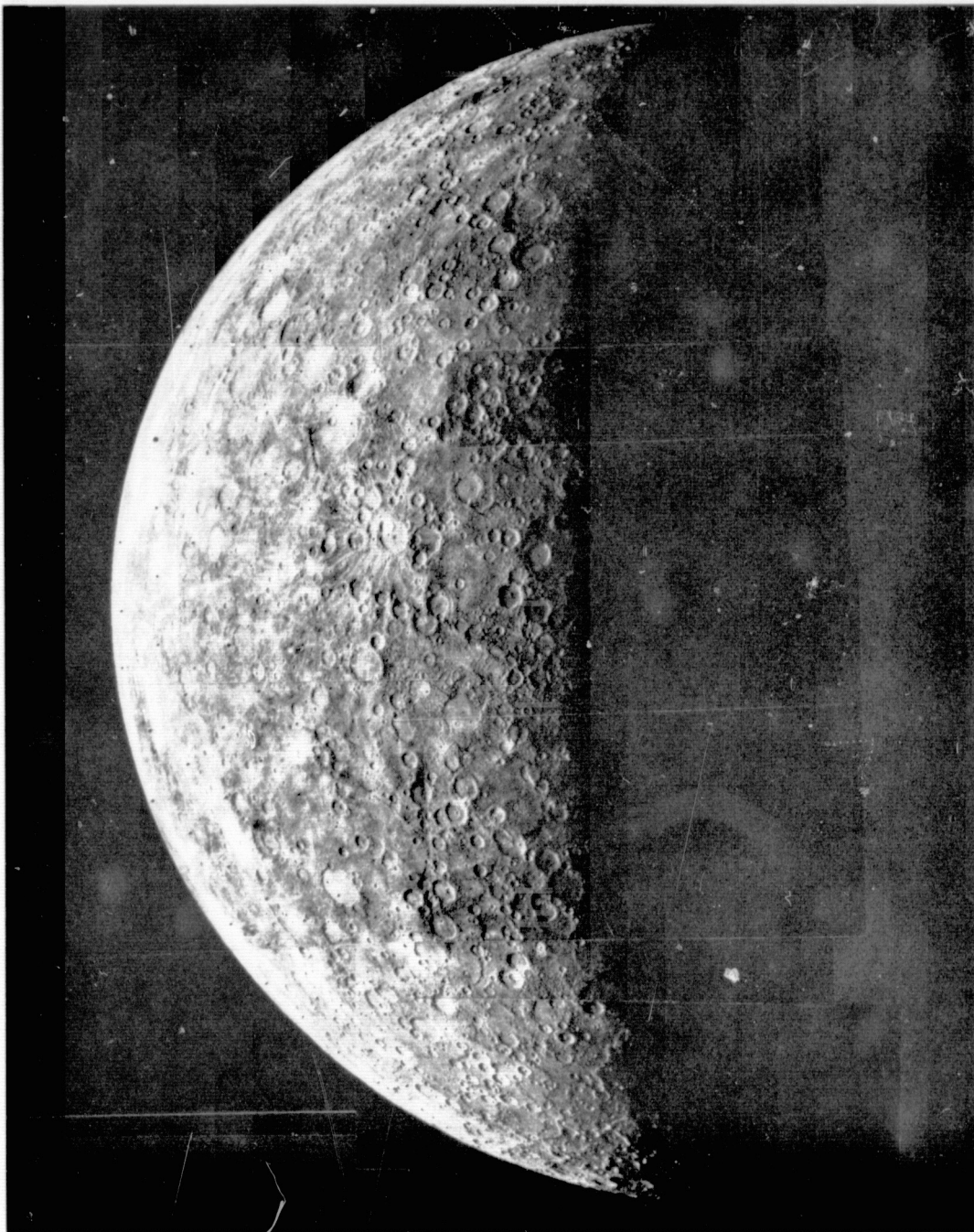
Figure 3. A photomosaic of low resolution UV and orange images of Mercury taken on the incoming leg of Mercury 1. These images were radiometrically decalibrated using appropriate decalibration conversion factors for each filter. The images were then high pass filtered and contrast enhanced; twenty-five percent of the DC component was included in order to retain the global appearance. (P-14470)
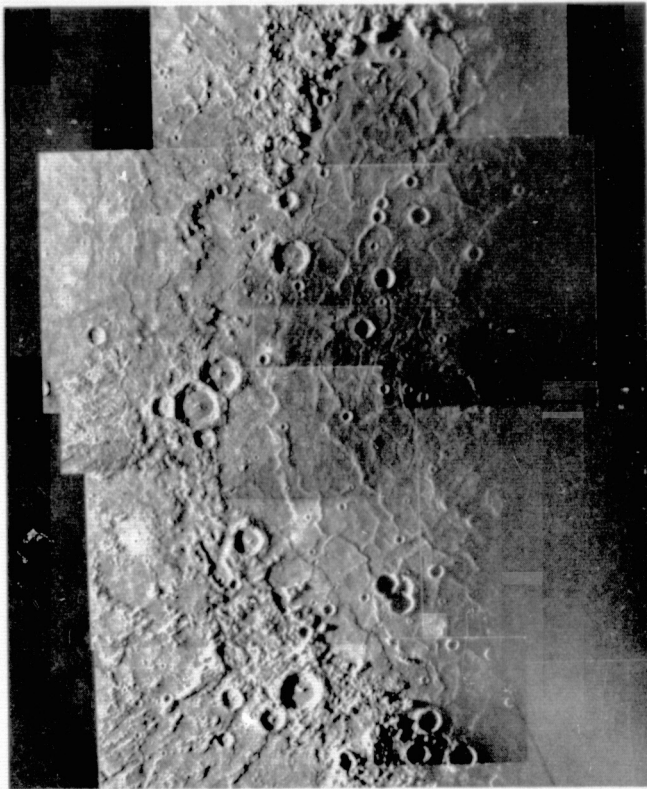
Figure 4a. A photomosaic of Mariner 10 pictures of the Caloris Basin of Mercury. These images were subjected to standard high pass filtering and linear contrast enhancement.
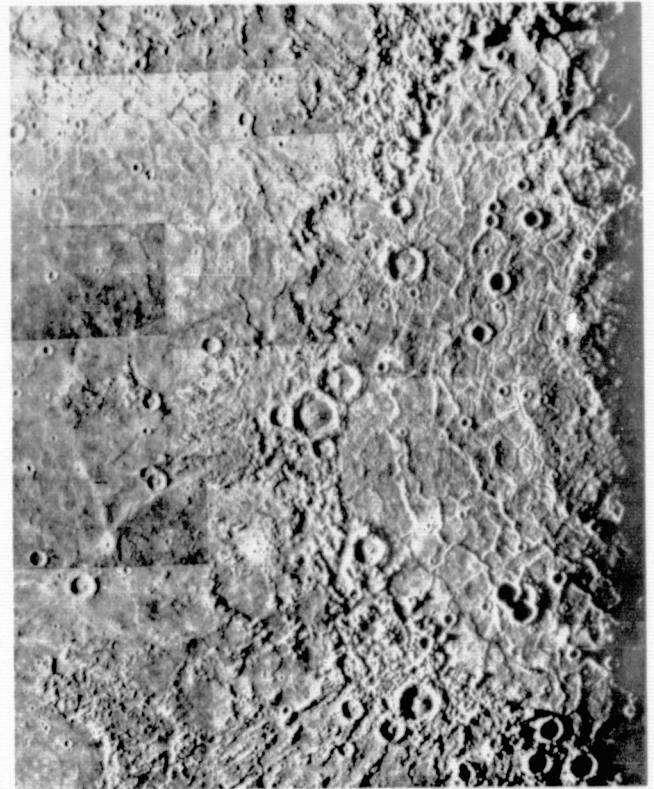
Figure 4b. A similar photomosaic except the filtering process included the scene dependent filter algorithm (Ref. 6). (P-14855)

Figure 5. A computer mosaic of an oblique stereographic projection of the Caloris Basin overlayed with latitude and longitude grid lines at 2 degree intervals. Processing of the images was similar to that of Fig. 4b.
(IPL Roll 749, PIC ID 75/7/28/175330)
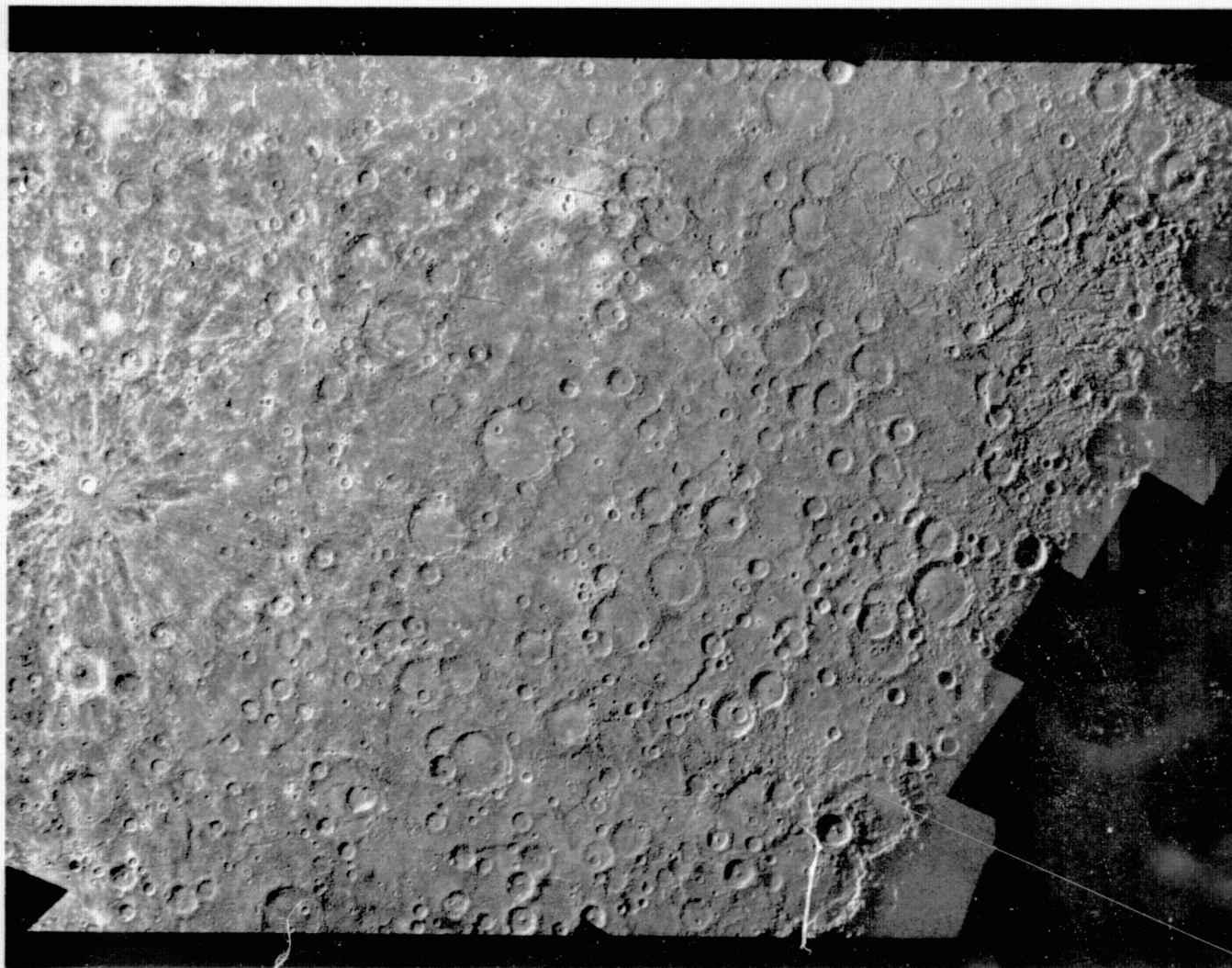
Figure 6. Computer mosaic of the H11 Quadrant of Mercury, covering the area latitude -70° to -20° and longitude 10°W to 70°W. This is a Lambert conformed projection composed of 76 Mariner 10 images. The output image is 7000x5000 pixels with a scale of 0.4 km/pixel.
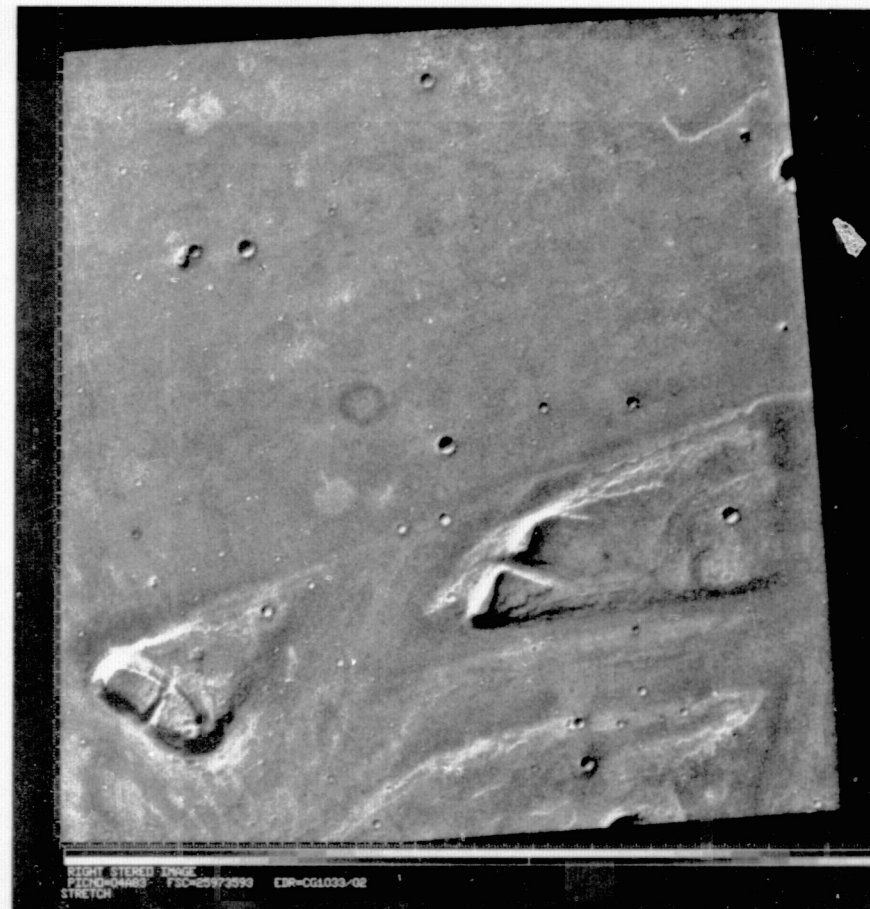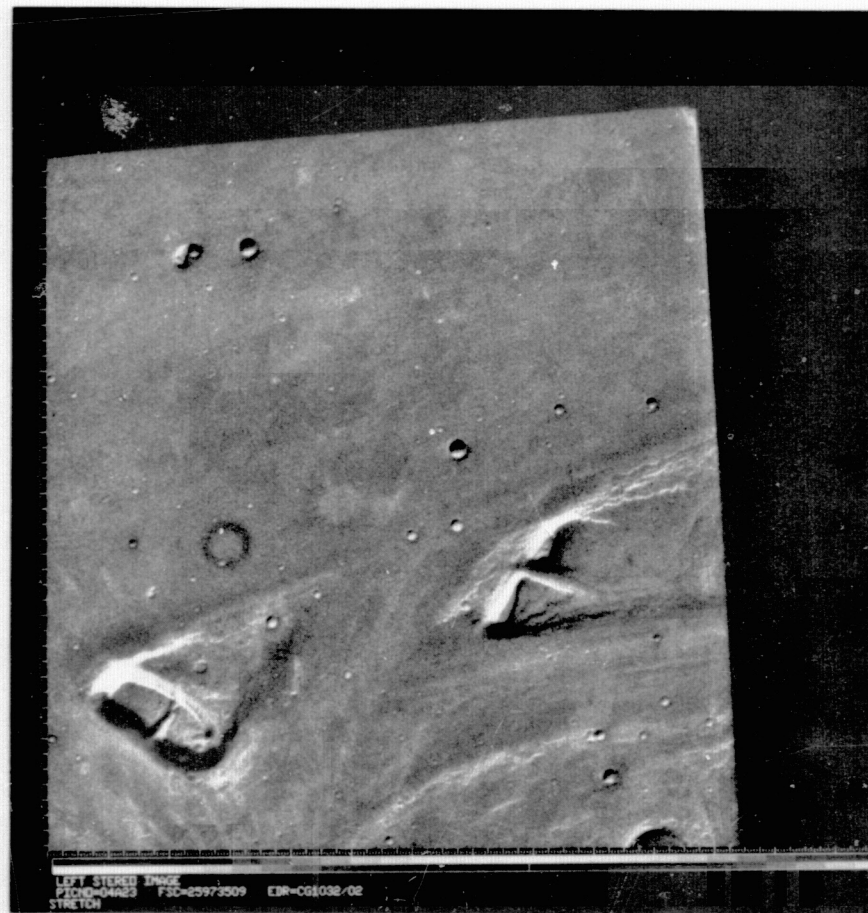
Figure 7. Left and right components of a stereo pair of images of Mars taken by Viking Orbiter 1.
(IPL PIC ID 76/09/28/164905 and 76/09/28/165232)
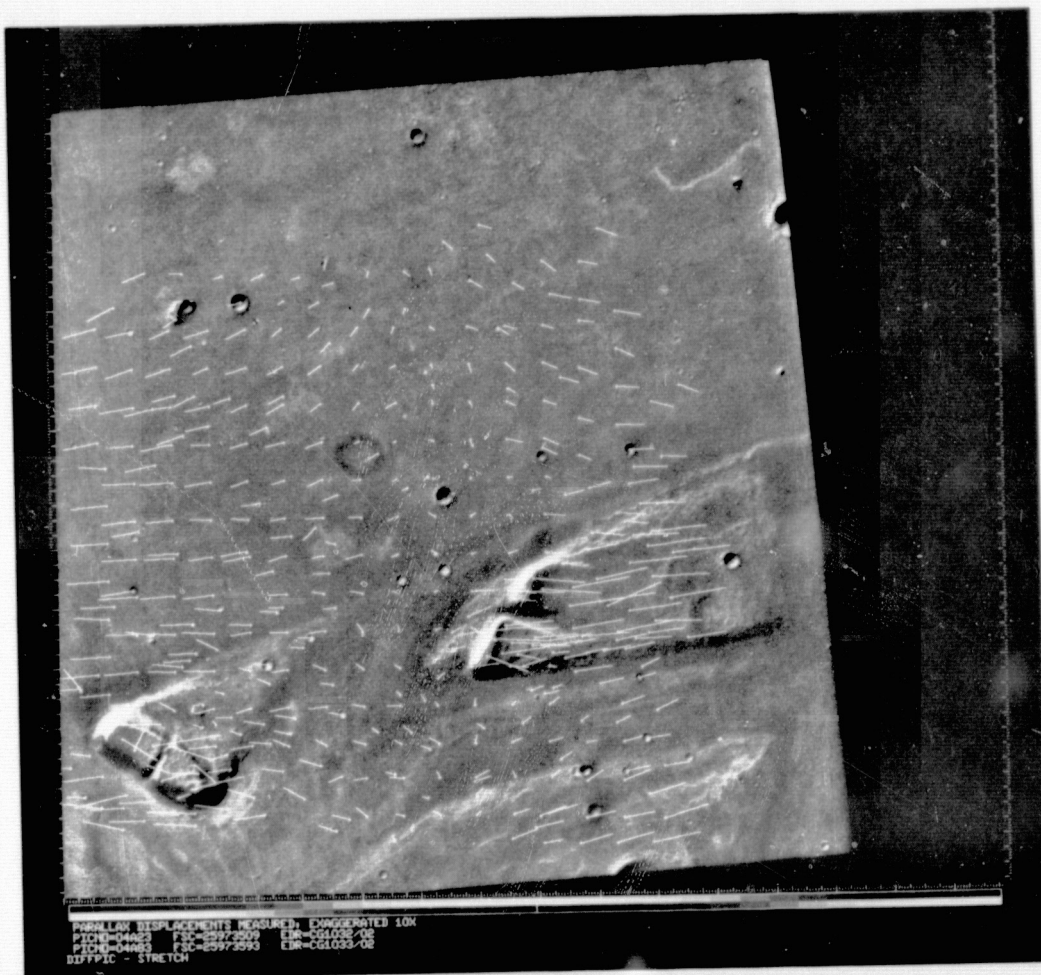
Figure 8. The right component of Figure 7 overlayed with displacement
vectors determined by the stereo photogrammetric program
STERGEN. The magnitude of the vectors is exaggerated 10X;
the origins are indicated by a dot.
(IPL PIC ID 76/09/28/165817)
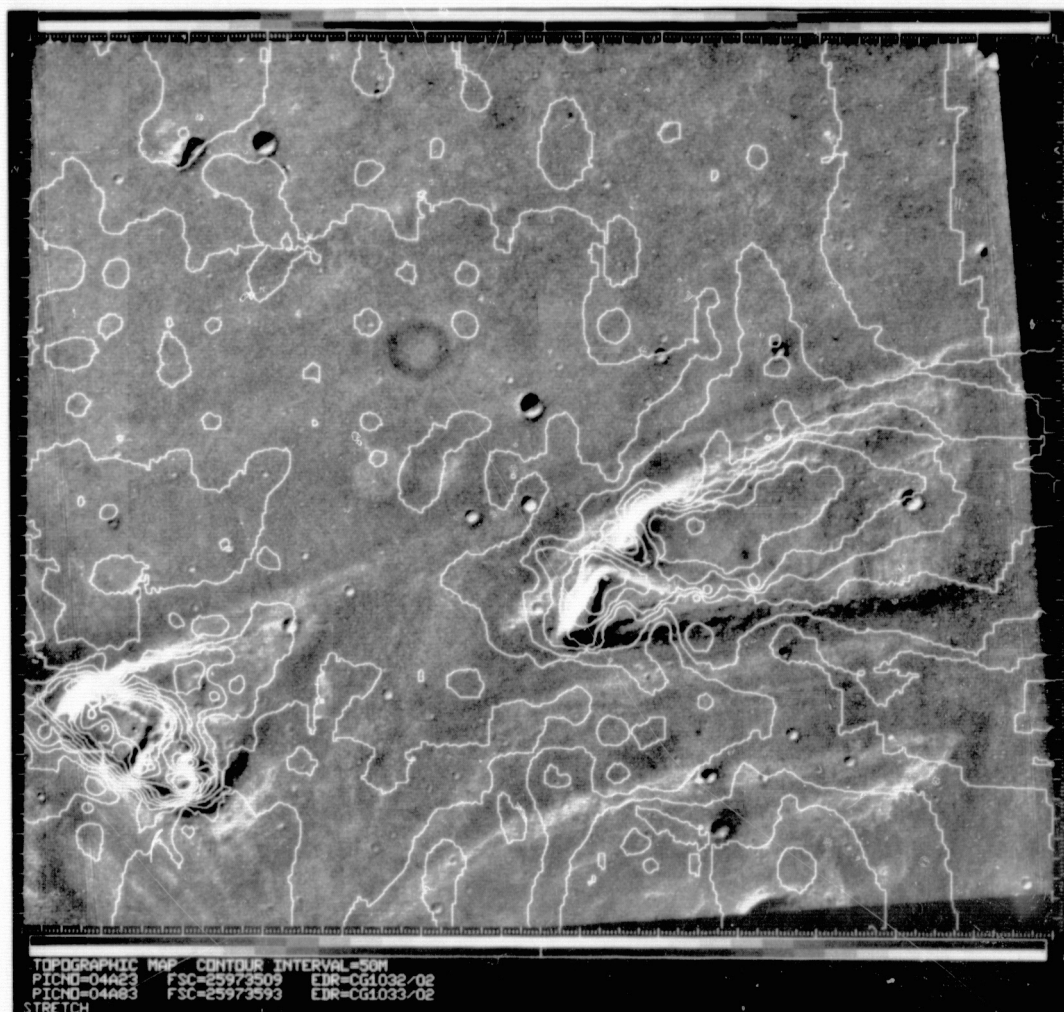
Figure 9. The right component overlayed with iso-elevation contour
lines. The elevation interval is 50 meters. Elevation
values were determined from interpolation between the vectoral
points shown in Figure 8. (IPL PIC ID 76/09/29/163437)
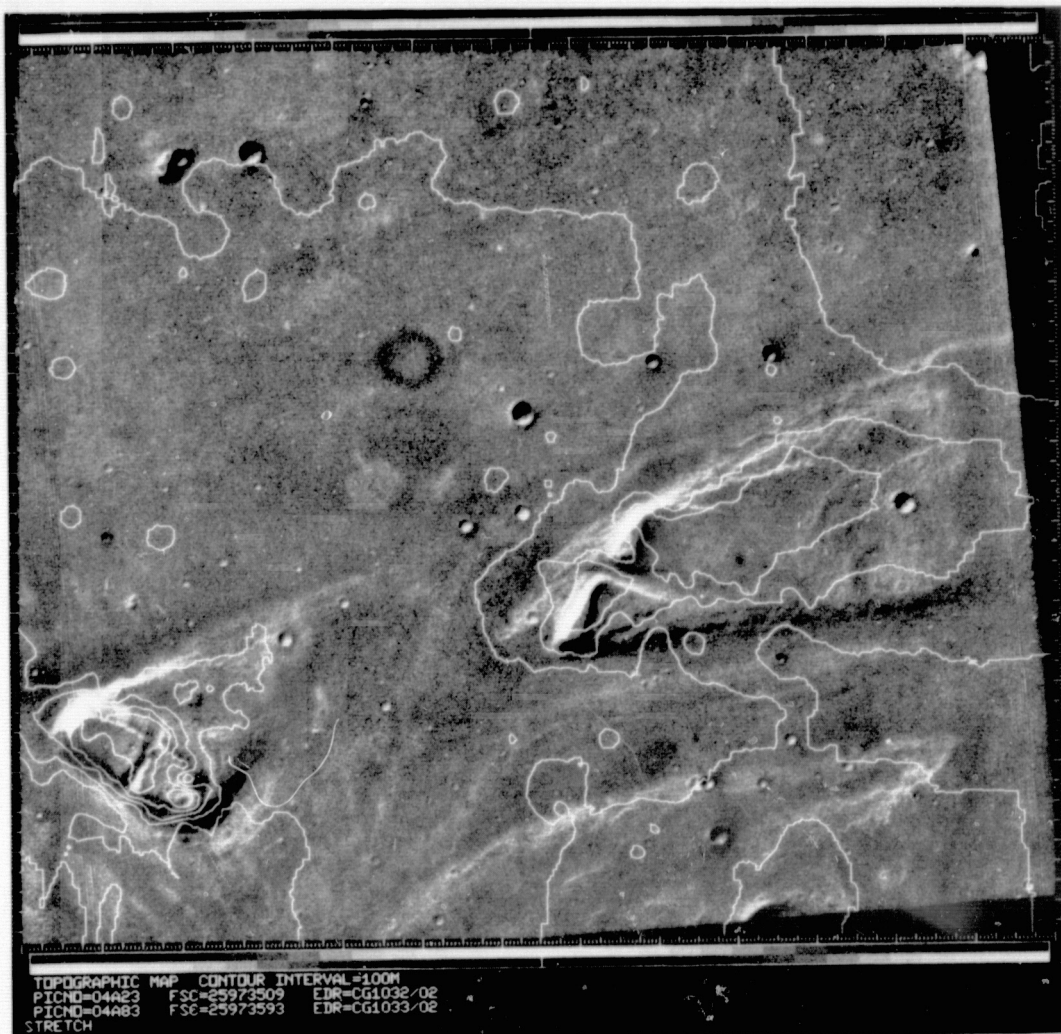
Figure 10. Same as Figure 9 except the iso-elevation contour interval is 100 meters. (IPL PIC ID 76/09/29/163012)
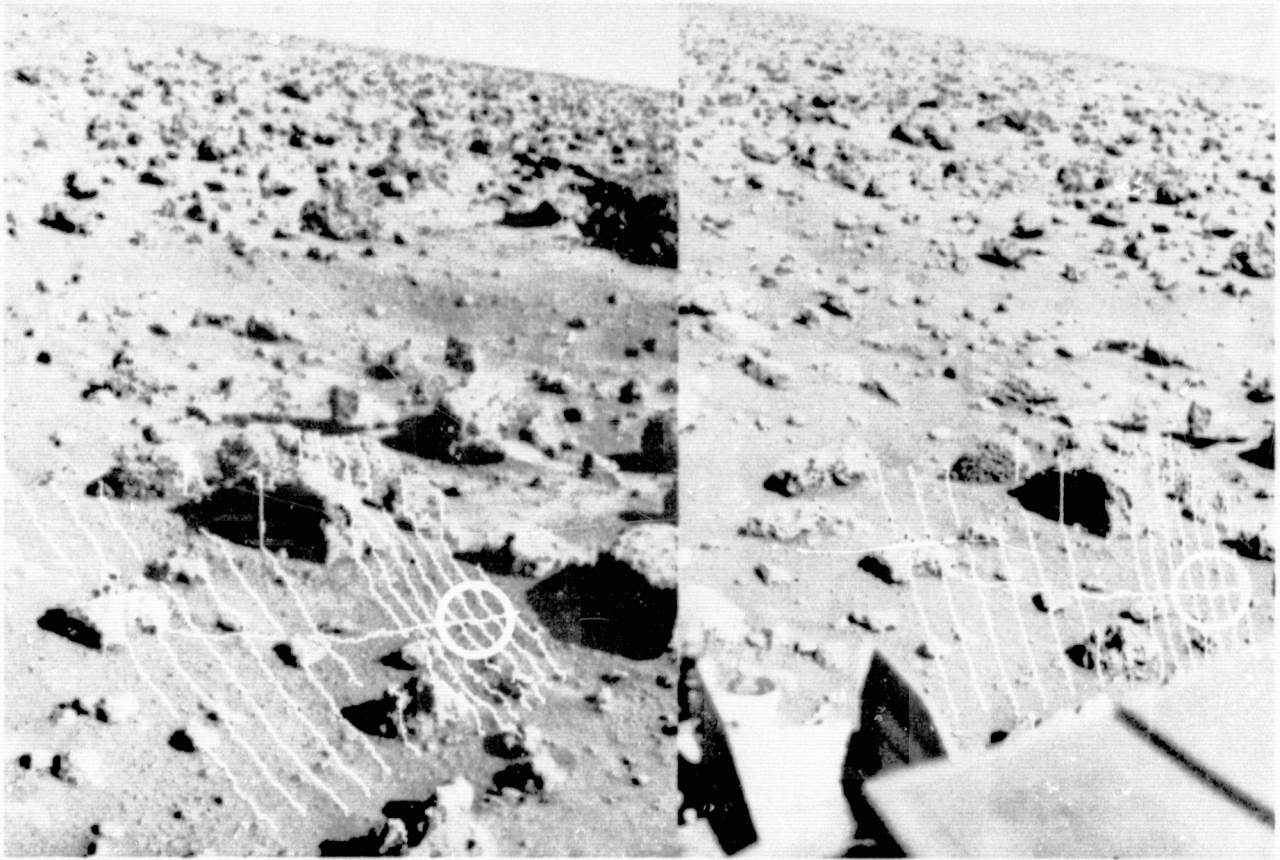
Figure 11.   Stereo images of a candidate surface sample acquisition site of Viking Lander 2. The white lines are traces of vertical profiles drawn along the surface in thirteen vertical planes, each of which passes through the surface sampler arm pivot point. The transverse profile is located in a plane at a constant radius from the sampler arm pivot point. The circle indicates the site which was chosen for the initial sample acquisition. These profiles were drawn using the interactive stereo photo-grammetric software and hardware system developed jointly by IPL and Stanford University. (P-17729)
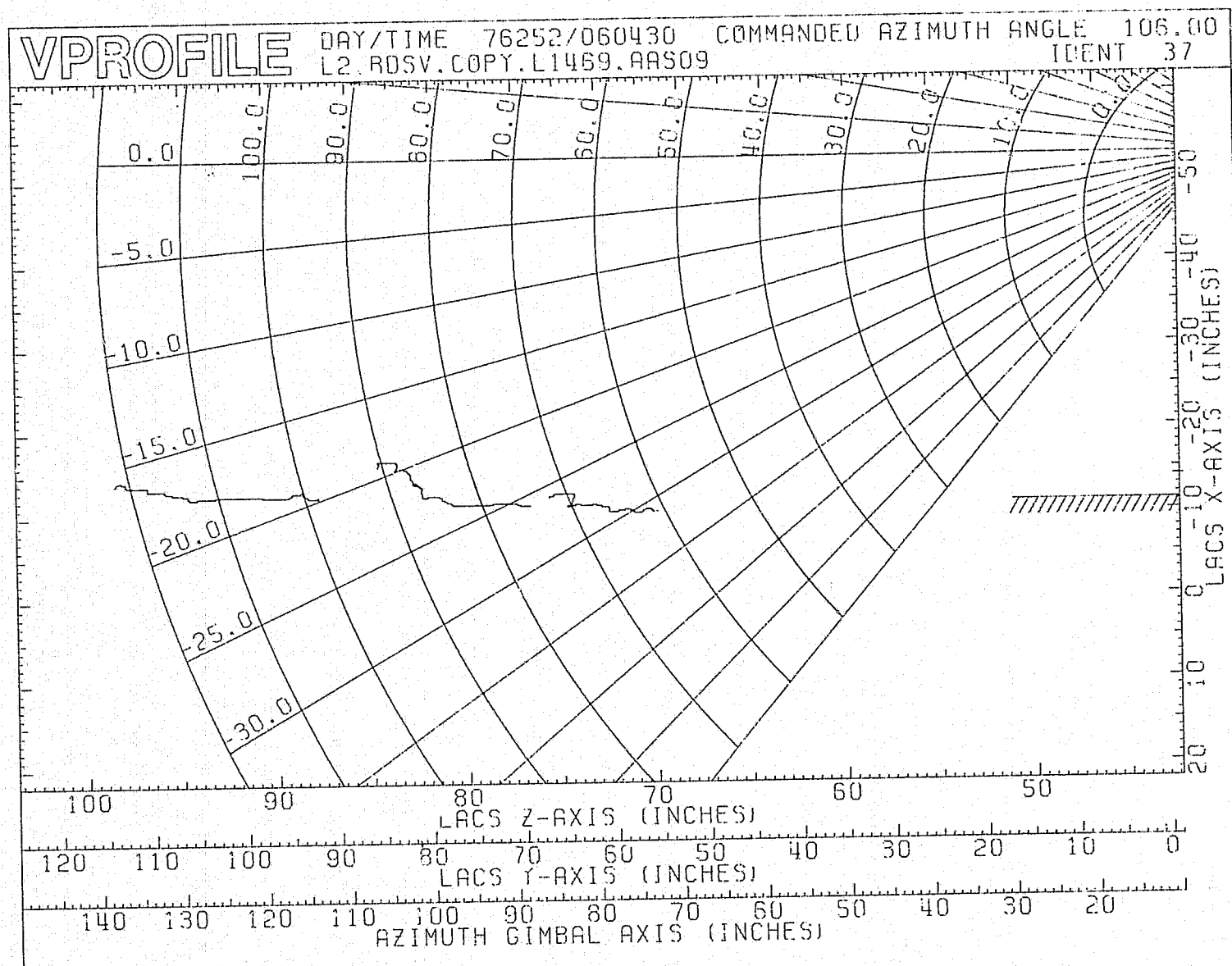
Figure 12. This is a VPROFILE plot of the leftmost profile indicated in Figure 11. The discontinuities are areas behind the rocks which are not visible in one of the stereo pairs. LACS coordinates refer to a rectangular coordinate system referenced to the Lander itself; X is up and down, +Z is forward and Y is left and right. The fan shaped grid indicates points of constant sample arm elevation and extension.

THE ROBOT'S EYES: VISION SYSTEM FOR THE JPL ROBOTICS PROJECT*

Donald S. Williams

Jet Propulsion Laboratory
Pasadena, California

ABSTRACT

The robot vision system, a part of the Robotics Research Project, is aimed at developing a description, for use by other subsystems, of the environment surrounding the television cameras and laser ranging device. The Robotics Research Project is developing techniques for control of semiautonomous vehicles and manipulators. Hardware preprocessing techniques are being explored to aid the image analysis routines (segmentation, region growing, and depth extraction) and to provide autonomic functions when the control computer is serving other subsystems. The images, from charge injection array cameras, are maintained in dynamic random access memory (RAM), which allows fast access by the analysis routines to any picture element. The analysis is executed on the robot control computer and a remote time-shared system.

I. ROBOTICS PROJECT

Why a robot? Two Surveyor spacecraft have carried out, on the Moon, tasks similar to those required to explore the Martian surface or the sea bed. The Russian Lunokhod vehicles were driven with some difficulty by controllers on Earth. These vehicles provided the Earth-based controllers, through a high-rate data link, with television pictures and engineering data. With only a three-second delay, at lunar distances, placing man in the control loop is a slow but an acceptable way to control the landers.

On the sea bed, a free-swimming vehicle linked to the surface by narrow-band sonar might require 8 minutes to transmit data. Between Earth and Mars the round-trip communication time varies from 6 to 44 minutes (Ref. 1), the data rate is limited to 16 kilobits per second on a direct link (8 kilobits per second via orbiter relay), and the communication window is limited. In this new situation, placing a ground controller in the loop is suitable only for simple tasks. Emergency response and accommodation to surrounding detail are no longer feasible under delayed direct control. It has been estimated (Ref. 2) that, under direct ground control and a round-trip communication time of 27 minutes, approximately 1.5 hours would be used for each 6 minutes of lander activity and which assumes an efficient ground operations (not 12 hours to 1 week turnaround). When the basic delays are compounded with intermittent communications due to planet rotation (Mars and Earth), relay satellite orbit, and multiple missions, extensive investigations under direct ground control are going to be lengthy. Thus, a robot.

At JPL, in a robotics research program sponsored by NASA's Office of Aeronautics and Space Technology, an effort is being made to demonstrate the feasibility of making a semiautonomous robot that will, on command, navigate safely from one place to another or manipulate nearby objects without direct ground control. Thus, the rover must interpret data that would, under direct control, have been transmitted to Earth for human interpretation. The semiautonomous rover would then generate an acceptable action sequence to complete the assigned task. The scientific objectives, mission strategies, target selection (including designation of "interesting" objects), and sequence plans would still be specified by the ground control. By combining limited autonomous operation with Earth control of the traditional lander functions, it has been estimated (Ref. 1) that the scientific return from a Mars surface exploration will be increased by a factor of 100 over a mission totally under direct control.

To carry out the demonstration, the robotics research program has been divided into subsystems: arm, vehicle, vision, ground system, and support (communications, computing, and executive software). While this is not the first autonomous vehicle project to combine mobility and manipulation with vision, this project seems to be happening when the necessary support technology is becoming available, particularly the economic parallel processing of significant capability and preprocessing of sensory data. This project is also concerned with the new requirements placed on the ground system by the semiautonomous operation. Since most robotics projects have dealt with either autonomous robots (no human control) or teleoperated robots (time delays generally of less than one-quarter second and rarely more than three seconds), the problem of delayed interaction and supervision has not been an issue. In the face of the usual conservative mission philosophy represented by "minimum change from the previous mission," a semiautonomous rover will be accepted only if the ground system provides a user confidence equal to or better than the traditional direct control system in addition to the increased scientific return.

II. VISION SYSTEM

There is no one correct way to sense the environment. Each subsystem has sensors adapted to the local problems encountered by the subsystem. But to avoid the problem of several blind men trying to describe an elephant or several subsystems describing, in their local terms, the Martian surface, a broader description of the environment is also required. In animals that description is provided by vision systems adapted to the animal's size, environment, and role as hunted or hunter. Animal vision systems may sample the environment both actively and passively in multiple-frequency bands, varying from audio through sonar up to frequencies around our visible spectrum.

The vision system for the robot provides a broad description by sensing the environment both passively and actively, using two solid-state television cameras and a laser ranger. Redundancy and cross-checking are provided since the cameras are sensitive to the infrared beam from the laser diode in the ranger. The two cameras can also operate as a stereo pair.

Machine processing of visual information is a complex area and thus not as well understood as, for example, speech recognition or chess. Each of these areas, representing a hierarchy of complexity, started with little preprocessing (abstraction) of the available data and large complex computer programs. For chess, the available data (piece location) is limited, has a high usable information content, and consists of, normally, only the present board (no history) position. In word recognition, the data rate has increased; much of the data can be discarded without loss of understanding and surrounding words (context or history) may be used to aid recognition. The data rate in vision systems represents an increase of another three to four orders of magnitude and more of the available data can be discarded. Today, a hand-held calculator is available which plays a simple game of chess, and computer chess tournaments are held regularly. Isolated word recognition systems are used commercially today, and much of the required preprocessing (abstraction) of the speech signal is done on one large-scale integration (LSI) chip. Vision systems are now showing the same development patterns as seen with chess and word recognition (Ref. 3). Researches in continuous speech understanding and scene understanding are developing areas; thus it is still too early to draw a similar parallel. In the cases cited, as an understanding (human, that is) of the data developed, improved abstractions of the data are generated, often in parallel, with a clear hierarchical structure with feedback being used for controlling the abstractions generated.

Today, the conference topic of imaging software is appropriate, but as the processing techniques become established more imaging hardware will be involved. Even today, as imaging software is discussed, the preprocessing done by hardware must be included. But, first, the transformations applied to images by the robot vision system are edge (transition) detection, region growing (based on gray scale or color), texture (frequency) analysis, ranging, and isolated segments. Isolated segment processing, used in both speech and image processing, accepts as correctly identified a region or edge that has a high probability of being correct and uses that segment to locate and identify related segments.

A. IMAGING HARDWARE

The present, largely software implemented, robot vision system represents the early stage of understanding of how to do image recognition. The vision system hardware consists of sensors, memory, displays, and the computer.

1. Sensors

The sensors are a pair of solid-state image arrays using a charge injection technique (Ref. 4) and a time-of-flight pulsed laser ranger. The

television cameras containing the image array and the laser assembly are mounted on a common pan-tilt assembly. The image array is 244 lines by 188 elements per line. Alternate lines (122) are output twice, once from a delay line, in order to generate a standard television field (244 active lines per field). The remaining lines are output in the same way to generate the second field. The cameras were modified by JPL to accommodate external sync, a larger variety of lenses, and critical adjustment while enclosed. The miniature power supplies furnished with the cameras are also targeted for replacement. Experiments with auto-iris lenses indicate that an alternate control strategy is needed to reduce wear as well as to provide a satisfactory image. The strategy will include control of lights on the vehicle. With a 25-mm lens, the cameras can resolve 5 mm at the normal working distance for the arm (2 meters).

The laser ranger uses a pulsed gallium-arsenide diode. The output power is intentionally kept low to insure safety. As a result, some light sources overpower the sensitive detector. The time of flight is measured with conventional nuclear instrumentation. The calibration of the laser ranger is checked frequently with reference points on the vehicle. The beam is steered by a two-degree of freedom flat mirror assembly based on the Surveyor imaging system. Resolution in the arm work area is 5.8 mm horizontally and 11.6 mm vertically.

2. Video Conversion and Memory

The bandwidth of a video signal and the available data volume greatly exceeds the capacity of conventional computers and their associated memories. Thus, an alternative to serial digitization at video rates is required. Only recently have video rate analog-to-digital converters become commercially available. The cameras (silicon target vidicons) originally used in the robotics project were operated at 729 lines per frame in order to meet a ranging accuracy requirement for long distance navigation. To handle the video data, a column digitizer was designed (Ref. 5) which sampled one point in every line during one frame time. Sampling a column in each frame provides sufficient time (one television line) for data transfer to the computer memory and writing blocks to disc storage. By advancing the column, an entire image could be converted. The column digitizer will operate at any line rate up to 729 lines per frame (60 fields per minute). Similar devices are now offered commercially for resolutions up to 525 lines per frame (Ref. 6). For higher resolutions special scanners are available. As the range accuracy requirement for long distance navigation was reduced, the need for improved geometric stability of the image became apparent, and increasing processing speed became an issue; thus a transition was made to the present solid-state array cameras and a video memory.

The new digitizer and dynamic memory (RAPID) operate at video rates (Ref. 7). The memory can be updated continuously, thus providing in effect a random access camera, or the memory can freeze a frame for analysis. The memory has interleaved synchronous ports for the video signals (both cameras and monitors) and asynchronous ports for computer access. The synchronous ports

are buffered in blocks of eight bytes to prevent memory access conflicts. Converters and memories suitable for full frame broadcast television signals are available commercially with either two video interfaces or one video and one computer interface (Refs. 8, 9, and 10). The digitizer in RAPID has been adapted to the eccentricities of the present solid-state image array so that redundant (and possibly inferior) data is not stored. The sampling window is adjusted to center on each sensing element in the array. The memory can store 256 elements by 256 lines from a selected camera or the computer. Since the memory capacity exceeds the image storage requirements, the remaining space is used for laser range data obtained by the computer. All of the memory is accessed during each frame so that a separate refresh cycle is not required. When an image is frozen in the memory, the memory is still accessed in the same pattern to refresh the dynamic memory. A television image is continuously generated from the memory for display. The image memory is interfaced to the computer input/output bus so that any point in the digitized image can be accessed in about 5 µs by specifying the eight-bit line and column address. The new access method is much faster than requests to a file handler, which must determine the storage location of the requested element (core or disc) and possibly retrieve the element from disc.

3. Displays

Three types of computer-controlled displays are available, each suited to different types of data. The raster display from the video memory has already been mentioned. It displays either the digitized video image or a computer-generated image. A second raster display (RAMTEK GX-100A) is used for processed image and laser data. This display has hardware vector and character generators and can only be written by the computer. For displaying the abstract data, a vector display (IMLAC PDS-1D) is available.

4. Preprocessing

As machine recognition research progresses in a specific field, preprocessing of the incoming data increases. The current limited preprocessing of video data indicates the present limited applicability of available techniques. The results and problems of image analysis using the present robot configuration suggest preprocessing techniques aimed at three phases of image analysis: reflexive actions, processing aids, and object detection.

The lens system is the first to get attention with work proceeding on automatic iris and focus, both reflexive actions. The problems of direct auto-iris have already been mentioned. Automatic focus, which shows variable performance due to scene dependencies and field depth, is necessary to maintain clear boundaries, and it also provides a rough distance measurement. Auto-focus has wear and reliability problems similar to those of the auto-iris mechanism though less severe. Slower acting controls with memory and input from both pan-tilt controls and current requirement for image data can eliminate many corrections and reduce others.

Processing aids are being investigated in four areas: foveal-peripheral vision, auto-zoom, edge enhancement, and object tracking. Foveal-

peripheral vision, provides high resolution in the image center to lower resolution at the edges, thus increasing coverage while maintaining high resolution at least in the center. Unfortunately, when several accurate position measurements are required, the camera may be repositioned many times, and the software must track objects (or parts of objects) from scene to scene.

Based on the area being analyzed, auto-zoom can adjust resolution and coverage. Due to the accuracy requirements placed on the output from the vision system, the repeatability of element positioning in a zoom lens is a problem. Graticule marks on the outer lens element are being considered but locating the points in the frame increases the software processing requirements.

Filtering, which is widely used as a preprocessing technique in the time domain for speech recognition, is more complex in imaging applications due to the increased dimensionality: color and spatial frequencies. The robot vision system is monochromatic, using either the full bandwidth of the cameras or the IR region alone. Spatial filtering has, thus far, been implicit in the segmentation and region growing algorithms. Filtering of the video signal in the time domain is attractive from a hardware viewpoint but of limited use since the filter is one-dimensional (horizontal), leaving the orthogonal (vertical) frequencies unchanged. Consideration is being given to the edge enhancement techniques used in television broadcasting, which are two-dimensional and similar to a characteristic of the human eye, as an aid to the edge detection software. Adaptation of the commercial equipment could provide addition spatial filter characteristics, particularly that of a low-pass filter. The low-frequency component can be subtracted from the original signal giving a signal which contains texture signatures.

Object tracking, such as used in "smart" bombs, could provide feedback in parallel with the motion of the arm or vehicle. The object tracker could be implemented in hardware to supplement the present minicomputer which was designed to control one function at a time or the tracker could be implemented in another processor which could be delegated all of the local robot image processing.

Image differencing was explored early in the development of the robot vision system for distance measurement. By differencing two pictures, one of which contains the laser spot, the range to the designated area can be determined, thus providing a check on the time-of-flight laser data. Image differencing is not currently used on the robot due to the processing time, if done in software, and the lack of equipment to do it electronically. Stereoptic fusion, by image differencing in a limited area, aids object detection and provides another range measurement. Presently, the capability does not exist, in the robotics project, to register the images either electronically or mechanically, thus leaving software registration which is too time consuming on the robot control minicomputer.

B. IMAGE SOFTWARE

The imaging software is aimed at two goals: object locating (arm) and transversibility (vehicle). Since responsibility for selecting interesting sites

will be retained by mission operations, the vision system is concerned with accurately locating objects in designated areas for manipulation and navigation. The vision system is not expected to recognize interesting objects, such as roman columns, since humans have not been able to provide a suitable representation to a machine on the concept of "interesting." That task is regretfully left to the ground system.

The framework for scene recognition, based on the output of several subsystems, is presently provided by a semantics network (Ref. 11), which is implemented on the DEC PDP-10 computer at Caltech. The remaining subsystems, scene segmentation (Ref. 12), region growing (Ref. 13), stereo correlation (Refs. 14 and 15), and camera calibration, are implemented on the robot control computer along with the arm and vehicle subsystems. The vision subsystems can provide inputs directly to the arm and vehicle subsystems.

1. Scene Segmentation

In a digitized image, drawing boundaries around objects to determine regions is a major step in analyzing a scene. In isolating such regions, it is usually assumed that an object has uniform local properties that may show a slow change across the surface and that detectable discontinuities in local properties occur at the edges between two objects. Under these two assumptions, the problem of object isolation is typically approached in one of two ways: edge detection and region growing. Region growing, to be discussed later, starts from a point or small region and finds all those points that belong, with a high probability, to the same region. Edge detection applies local independent operators to small regions throughout the image to detect and possibly follow the edges. Edges, unfortunately, are frequently fuzzy, variable in character, and discontinuous. The scene segmentation algorithm developed for the robotics project is fast, executes in a minicomputer, and handles multiple objects.

The basic structure consists of an edge detector which inserts boundaries even in regions of low probability and a region melt which combines similar neighbors. Edges that do not terminate or terminate without enclosing a region (cracks) could be handled by a path generator and special region grower. The special region grower would form two regions with the crack as a part of a common boundary. The path generator would extend a boundary using a shortest path to another boundary. The sensitivity of the edge detector has been set high enough to avoid the need for these two routines.

In edge detection, non-intersecting neighborhoods around the points in question must be used due to noise (both system and from the original scene). Thus, the neighborhoods must be properly selected and appropriate measurements made in order to exactly position the edge. To measure the possibility of an edge between two neighborhoods, values are taken in a region normal to the proposed edge. Maximum likelihood estimates are calculated assuming the values came from one object or two objects. The ratios of the two estimates, which is stored in a matrix for later evaluation, determine which assumption to select. Presently, four neighborhoods are defined to deal with edges, closely spaced lines, and corners. The maximum of the maximum likelihood ratios

for the four neighborhoods is used for the ratios. A separate matrix of ratios is determined for horizontal and vertical edges. To locate the actual edges, minimums are located in the likelihood ratio matrix after setting all values below a predetermined threshold to zero. Regions are expanded from the minimums as far as possible without crossing a maximum value. The resulting regions and edges are cleaned up by combining small regions with their closest neighbor and deleting all cracks. The entire process was accelerated by checking the region around the point in question to determine, based on local changes in intensity, if an edge might be present. If no edge is indicated, the output to the matrix is zero, and the maximum likelihood estimates are not computed.

2. Region Grower

The region grower starts at a designated location and, using local statistics, determines if neighboring points belong to the region. Three different criteria are available for determining membership in the region: absolute limits, local statistics, and extrapolation.

In the first option, points are added to the region until a specified range of intensity values is reached; then neighboring points are added only if their intensity is inside the existing range. Growth stops when no adjacent points can be added.

The second option uses a threshold matrix and two statistics: average intensity for member points and the variance, both weighted by the most recently added points. A test value,

$$\frac{current\ intensity\ -\ weighted\ average,}{variance}$$

is compared to the first pair of entries in the threshold matrix. If the test value is below the lower value, accept the point. If the test value is above the higher value, reject the point. Otherwise, try a neighboring point. The test value for the neighboring point is compared to another pair of thresholds. The test continues until a point is rejected or accepted. Since the last threshold pair in the table is set equal, a decision is guaranteed and the maximum search length is preset.

Option three replaces the average intensity in option two with an expected value which is calculated for the present coordinate position. The expected value is based on the best plane fit through the values of the points already belonging to the region. Nine statistics are calculated for each point. Options two and three are about equal in overall performance but each does better than the other in some situations.

3. Depth Extraction

In developing a map which accurately mirrors to sufficient detail the original scene for manipulation and navigation, range data is significant in determining the terrain characteristics. Range data is provided by processing the dual camera images and the time-of-flight laser. One serious problem is correlating laser range data with boundaries detected in the images since laser-detected changes do not necessarily match changes in light level detected by the cameras. Deriving range data from the camera images removes the

correlation problem. The laser spot can be detected in the camera images, providing an alternate measurement. Two methods have been developed to extract range data from the camera images: adaptive window (Ref. 15) and region matching (Ref. 14). In either case, the upper and lower bounds on the object distance from the two cameras can be established on prior knowledge. The bounds reduce the search space. Since only the distance to objects and their boundaries are necessary, a complete range map is unnecessary, thus providing another way to reduce the search space.

In using cross-correlation between two images to extract range data, tying all points which are visible in both scenes together is unnecessary if not impossible. Contour lines over most of the surface are predictable and thus are not interesting, only the changes represent new information. Natural changes occur usually where texture is rough as opposed to areas of smooth texture. Thus, the adaptive correlation window is used to limit searches where the texture is smooth (low variance) or to increase the number of tie points (located corresponding points) where the surface texture is rough (high variance). Even with those restrictions, the change from one tie point to the next tends to be small, thus allowing a reduction in the window size (the number of points used for a correlation) and search space. The window size is increased only when the correlation threshold is not reached. This technique was originally demonstrated using photographs from a dry river bed near the laboratory and was also used to process stereo planetary imagery.

As the scene segmentation and region growing algorithms were improved, depth information was extracted by comparing common regions in the two images. The region matrix contains far fewer points than the original image, and the boundaries are of prime interest anyway. If the existing knowledge of the possible ranges is used in a given area of an image, then the search for correspondence is again limited. The present algorithm requires a close match between the two regions before accepting the two regions as a match. That match requires high correlation in the two views and in the segmentation. It may be appropriate to map one object into another object rather than mapping onto another object.

III. CONCLUSION

Imaging processing is a developing field starting from handling small two-dimensional image processing to large two-dimensional scene processing and analysis, and now to three-dimensional analysis. The progress has depended on an improved understanding of the problem and increasing capabilities of the available techniques, particularly in reducing software processing time. Much of the process efficiency depends on what is done early in the process.

The robot's eyes now include a hardware configuration designed for fast software processing of the data. As experience is gained with the problems of robot control, additional hardware that is tailored to those problems will be explored. The applications include reflex action and aids to the higher level image analysis software. Some of the

proposed hardware will allow vision functions to proceed while other software systems are active.

The basic vision functions of rapid isolation of objects and ranging for manipulation and navigation have been demonstrated. Now a fast, reliable integrated system for semiautonomous robot directions is needed.

REFERENCES

1. Hooke, A. A., Larman, B. T., and Whitney, W. M., "The Impact of Robots on Planetary Mission Operations," International Telemetering Conference, Session on Robotics and Telemetry, Los Angeles, CA, October 1974.

2. Moore, J. W., Hornbrook, G. K., McDaniel, W. S., Gilden, J. R., Dorrah, W. E., Swirdling, M., Bank, H., Imus, R. E., Gottlieb, T., Lim, L. Y., Kurz, D. W., and Roberts, P. H., "An Exploratory Investigation of a 1979 Mars Roving Vehicle Mission," Document 760-58, Jet Propulsion Laboratory, Pasadena, CA, 1970 (JPL internal document).

3. White, G. M., "Speech Recognition: A Tutorial Overview," IEEE Computer, Vol. 9, No. 5, pp. 40-53, May 1976.

4. Model TW 2000 Closed Circuit Television Camera, manufactured by Optoelectronic Systems Operation, General Electric Company, Electronics Park 3-201, Syracuse, NY.

5. Alsberg, H., and Said, N., "Column Digitizer: An Interface Between a Stereo TV Camera and a Computer," Document 750-47, Jet Propulsion Laboratory, Pasadena, CA, April 1974 (JPL internal document).

6. Image Scanner, manufactured by Spatial Data Systems, Goleta, CA 93017.

7. Yakimovsky, Y., Rayfield, M., and Eskenazi, R., RAPID — A Random Access Picture Digitizer, Display, and Memory System, Technical Memorandum 33-772, Jet Propulsion Laboratory, Pasadena, CA, May 1976.

8. Matley, J. B., "A Digital Framestore Synchronizer, "SMPTE Journal, Vol. 85, No. 6, pp. 385-388, June 1976.

9. Frame Grabber, manufactured by RAMTEK Corp., Sunnyvale, CA, 94086.

10. RCA Broadcast Systems, "Digital Video Synchronizer," SMPTE Journal, Vol. 85, No. 6, p. 442, June 1976.

11. Yakimovsky, Y., and Cunningham, R., DABI — A Data Base for Image Analysis With Nondeterministic Inference Capability, Technical Memorandum 33-773, Jet Propulsion Laboratory, Pasadena, CA, May 1976.

12. Yakimovsky, Y., Boundary and Object Detection in Real World Images, Technical Memorandum 33-709, Jet Propulsion Laboratory, Pasadena, CA, November 1974.

13. Yakimovsky, Y., and Cunningham, R., On the Problem of Embedding Picture Elements in Regions, Technical Memorandum 33-774, Jet Propulsion Laboratory, Pasadena, CA, June 1976.

14. Yakimovsky, Y., "Extracting Depth Information from a Stereo Pair," Milwaukee Symposium on Automatic Control, pp. 311-316, 1974. Also, Yakimovsky, Y., and Cunningham, R., A System for Extracting 3-Dimensional Measurements From a Stereo Pair of TV Cameras, Technical Memorandum 33-769, Jet Propulsion Laboratory, Pasadena, CA, May 1976.

15. Levine, M. D., O'Handley, D. A., and Yagi, G. M., "Computer Determination of Depth Maps," Computer Graphics and Image Processing, 1974.

PANEL ABSTRACTS

"CURRENT FRONTIERS IN IMAGE PROCESSING TECHNOLOGY"

F. C. Billingsley, Chairman

Participants:

Patrick E. Mantey

James R. Lucas

David A. Landgrebe

Harry C. Andrews

Samuel S. Rifman

Andrew G. Tescher

THE BOTTOM LINE IS UTILIZATION

Fred C. Billingsley[*]

Jet Propulsion Laboratory
Pasadena, California

ABSTRACT

The lunar and planetary exploration missions — Surveyor, the Mariners, and Viking — have returned vast quantities of images to earth. In addition to the science processing of the images, operation of the missions has required the processing of the images in something approaching real time to allow operations decisions to be made as a result of them. This has required in turn the careful and efficient design of the digital image processing system. However, these missions are essentially "one shot" in that they each have a finite lifetime and a foreseeable end to the data flood. The Earth Observations satellites — Landsats 1 and 2 (now operating), Landsat C and Seasat (1977), Landsat D (1981) — on the other hand, promise to be forerunners of a relatively continuous stream of image producers with data rates in the >100 megabit/sec range.

The receint, preprocessing, and analysis of these images has been likened to trying to drink from a fire hose. Since this type of data processing pressure is driven by, or at least characterized by, the "Earth Observations Problem," let me set the stage today in terms of these satellites.

First of all, when considering the utilization of this data, we acknowledge the two types of analysis: (1) R&D, in which the images are exhaustively processed and in which processing time is a secondary consideration, and (2) "Operational," characterized usually by the need to process quickly and economically, and often to relate the images to other data. These two uses are not independent of each other, as the rapid and economical processing will come only as a result of previous R&D, and since R&D often utilizes images obtained during real operations.

In the early 80's we may well be faced by the following situation. Two Landsats in operation, each delivering perhaps 50 scenes of 6000x6000 pixels in each of six spectral bands, plus matching 3240x2340 four-band images, plus a Seaset delivering synthetic aperture radar data. There may be 100 users holding their breath waiting for the images (plus others in the wings), some requiring the images in one or two days, but with the mean delay perhaps in the three to four week range. Most of these users will relate these images to previous images, to maps, or to other nominal data. These users have committed themselves to utilizing the satellite data to solve their problems, so that data continuity and delay time for delivery to them is critical.

Now, what do we do to plan for this situation? We are immediately faced with a barrage of questions:

● Brute force gathering of the data seems not to be a problem. But are we gathering the optimum data? Can we be more clever in the gathering to pack more information into the data than by simple scanning? What is needed to design the data stream to allow maximum exploitation of the data?

● How smart do we make the sensor? Are there spacecraft considerations (e.g., attitude control) that will minimize the data handling load?

● How many different products are necessary? Which ones may be only niceties?

● How far downstream do we go in the processing before delivery to the user? Should we do all the radiometric and geometric corrections, or leave these to the users? If we do much preprocessing, have we caused some users the need to undo our efforts before they can use the data?

● Many of the users will use the images related to geographical location. Do we do anything at all, simply supply them with geographical pass points, or provide images completely registered to some standard cartographic projection? How about supplying images oriented E-W to simplify their data extractions?

● Do we provide any pre-analyzed images — for example, land cover homograms?

● Many users will require access to previous images, perhaps as early as the '72 Landsat images. How do we plan a data bank for this? Do we utilize data compression for data storage? How do we determine which images to discard?

● Should we develop . ' have available designs for data base systems su table for use with images?

● Do we have to actively pursue (i.e., fund) development of image analysis technology, or rely on the users to do so? How about providing actual software for fundamental processes such as data tape reformatting?

● What will be the hardware state of the art by then? What need we to do to be sure it will be what we need?

● The present college students will be users by the time these satellites are operational. Do we attempt any training at the college level. If so, what?

Although some of these questions seem purely policy, their answers critically affect the drive for technology development.

---
[*] On temporary assignment with NASA Headquarters in Washington, DC.

IMAGE-DERIVED DATA AND APPLICATIONS IN PLANNING

Patrick E. Mantey

IBM Research Laboratory
San Jose, California 95193

ABSTRACT

The data developed from satellites or other remote sensors represents an additional source of information that can be of significant value in planning and resource allocation. To realize the value offered by this data, it must be integrated into a data base so that it can be related to data derived from a variety of other (terrestrial) sources. To be used by decision-makers, this integrated data base must be supported with (interactive) retrieval and display functions that will make it readily accessible to (non-programmer) planners and managers. Today, agencies that are responsible for these planning and resource management functions are not prepared to support these uses of the data.

SUMMARY

In applications such as regional planning or resource management of recreation, agricultural or forest lands, data derived via remote sensing can provide information on land use, vegetation, rainfall, and other important variables. Other relevant sources of data in these applications include the census, tax assessment records, engineering files of surveys and road projects, building permits, utility connections, etc. An integrated data base developed from these diverse sources is required in most applications; i.e., rarely does any single source provide all the required data. Data extraction (Ref. 1) is defined as the process by which an integrated subset of data is developed from the source files relevant to a particular management or planning application. Extraction thus includes data reduction and file integration. Data reduction is required because:

a. the potentially useful data base will be much larger than the data actually used,

b. the user will want access to varying levels of detail in the data base,

c. the relevant subset of data will vary during the problem-solving process,

d. some data may not be compatible at the detail level of the data captured in the source files.

Extraction thus provides the user with a "virtually" integrated data base tailored to his application, without requiring total integration of source files into a common data base; and also offers possible performance advantages (Ref. 2).

In the planning and resource management applications mentioned previously, relating of the diverse data sources can be achieved by the common (sometimes implicit) references to geography, using a computerized Geographic Base File (GBF). Functionally, the GBF is a computerized file containing data on geographical features (e.g., rivers, roads), on buildings (coordinates or address ranges on street segments in cities), administrative overlays (e.g., census tracts, counties, planning areas) so

that a computerized map can be displayed of any of these features. Depending upon the detail level of the GBF and of the data and geographical references in the source files, data from these files can be integrated into subsets at the required level of detail to support a particular application. The challenge in developing and utilizing these integrated data bases today are:

1) GBF's are not universally available, and where available are often inaccurate and not well maintained.

2) Geo references are not contained in source files at the detail level needed by applications (e.g., tax assessor files with only APN and owner address, not situs address or centroids.

3) Continuing sources of data derived from remote sensors and compatible with administrative overlays (i.e., zones) appropriate to a particular application are not available.

4) Much relevant data available is only available for predetermined zones which are not compatible except at aggregations too coarse to be of much value.

5) System and staff support in most agencies is inadequate to support transformation of available data into relevant information for a particular decision.

6) Data management systems available today do not adequately support the storage and retrieval of data in both image and computer encoded forms.

7) Graphics hardware is not widely available to support the interpretation and analysis of geographically-related data.

Until all of these challenges are met, the rich source of data offered by remote sensors cannot be utilized commensurate with its potential by agencies reponsible for planning and resource management.

REFERENCES

1. Mantey, P. E., and Carlson, E. D., "Integrated Data Bases for Municipal Decision-making." AFIPS Conference Proceedings, Vol. 44, AFIPS Press, Montvale, NJ, 1975, pp. 487-493.

2. Carlson, E. D., "Using Large Data Bases for Interactive Problem Solving." IBM Research Report RJ 1585. IBM Research Division, San Jose, California, May 1975.

COMMENTS ON LAND CLASSIFICATION OF SOUTH-CENTRAL IOWA

FROM COMPUTER IMAGERY

James R. Lucas

Iowa Geological Survey
Iowa City, Iowa

ABSTRACT

Because the photographic laboratory at the Jet Propulsion Laboratory found it difficult to guess what were the optimal color renditions of the Iowa landscape for land classification purposes, the Iowa Geological Survey developed its own in-house capability for producing color products from digitally enhanced Landsat data. Research has now shown that efficient production of enhanced images requires full utilization of both computer and photographic enhancement procedures. Experimental photo-optical enhancement techniques carried out at the EROS Data Center (EDC) for a Landsat-1 29 August 1972 scene demonstrated that this imagery was more easily interpreted for land classification purposes than the EDC standard color composite. If digital processing procedures can be minimized by the utilization of photographic enhancement techniques, then costs for enhanced products could be reduced greatly in terms of both time and money.

Perhaps one of the major objectives of computer processing of Landsat data should be to produce high quality cosmetically enhanced data (eliminate striping or line drops) which has a range of digital values that can be recorded completely by film media. If digital processing can produce a digital distribution of data which fills the entire recording range of photographic film, then photographic enhancement techniques could be employed inexpensively by serious users familiar with the type of information they wish to extract from the imagery. Alternatively, if custom enhanced imagery is to be produced for users not having a photographic capability, different enhancements may be tried in an iterative fashion in the photographic laboratory, without having to preprocess digital data in the computer. The savings in time and money should be considerable with this approach.

A PERSPECTIVE ON EARTH RESOURCES IMAGE DATA PROCESSING

David Landgrebe

Purdue University
West Lafayette, Indiana

ABSTRACT

The pacing item for development of multispectral image processing in the earth resources field is data availability. The major source of data for this approach in the last four years has been the Landsat satellites which, spectrally speaking, produce rather primitive data. As we look forward to multispectral scanners with additional spectral bands and thus more complex data, correspondingly more complex processing algorithms will prove to

be desirable. Additional problems requiring work are 1) the incorporation of image spatial and temporal features into analysis algorithms, 2) cost effective implementations of existing algorithms, 3) more efficient and effective use of the human in processing operations, and 4) well engineered output formats which are inexpensive and responsive to user needs.

IMAGE PROCESSING AT THE UNIVERSITY OF SOUTHERN CALIFORNIA[*]

Dr. Harry C. Andrews
University of Southern California
Los Angeles, California 90007

The Image Processing Institute at the University of Southern California has been in existence for approximately the past 6 years. The Institute has been involved in a variety of application areas utilizing both optical and digital signal processing techniques applied to various image processing situations. While early results from research at the Image Processing Institute were based upon bandwidth reduction attempts and image restoration applications as well as image enhancement procedures, subsequent direction has taken a turn towards three topic areas that I wish to speak about today. These three areas are known as a) image understanding b) image restoration and c) mathematical problems for image analysis. While it would be presumptuous of me to suggest that this is the only research underway at the Image Processing Institute, these three categories are fairly representative of some of the topics now under investigation and are the three that I choose to address here.

Image Understanding is a new direction provided us by the Department of Defense through the Advanced Research Project Agency and represents a new approach to the problem of marrying computers to imaging systems. Image Understanding is really directed at learning, automatically, information from an image without the aid of human intervention. In its most grandiose interpretation one might consider the objective of an image understanding system to be capable of feeding an image into a machine and obtaining at the output of the machine a report written in English describing what is on the image in terms that are understandable and informative to a human reading the report. An example might be the feeding in of an aerial photograph into a machine and the report coming out describing wheat fields, urban areas, oceans, clouds, deserts, mountains, etc. While the image understanding objectives might first appear grandiose, it is my opinion that ARPA has undertaken a fairly sensible approach to a very complex problem. They have outlined a marriage of signal processing tools and artificial intelligence techniques to bring to bear upon the understanding task. Traditionally USC's expertise has been in the area of digital image signal processing. We are now gearing up to increase our staff in the area of artificial intelligence and computer vision personnel to tackle the image understanding objectives. One of the immediate objectives to which we feel we can contribute in this image understanding task, is the development of smart sensors. Smart sensors mean different things to different people. Probably the simplest example of the smart sensor would be a device which at the sensor did some sort of signal processing which lent its output to an understanding phase. Thus, one might build a sensor to measure texture, to measure edges, and to possibly automatically segment an image prior to any interpretation or subsequent transmission. We are actively engaged in the direction of the construction of a CCD signal processing array to indeed implement such smart sensor devices. The tools and techniques that I foresee as being quite useful that we will borrow from the computer science field and the artificial intelligence repertoire will include the use of list structures, languages and grammar and their various relationships to make a heterarchial image understanding systems configuration. Probably one of the more significant efforts will be in the development of the use of a world knowledge base to aid in the addition of contextual information to a particular imaging situation. By this is meant that if we have a sensor, contextual knowledge will tell us what the sensor is looking at in general. If the sensor is looking at an ocean, one would expect that the software and structure for land mass understanding would be switched out and the software and structure for ocean phenomena understanding would be switched in. Various additional examples of the use of artificial intelligence tools applied to image understanding could be developed but will be postponed for oral discussion.

The topic of image restoration is one that has had some interest at USC, probably because of the unique opportunity of applying some fairly sophisticated numerical analysis procedures to two dimensional image processing. You may have heard or been aware of some of the singular value decomposition approaches resulting in the pseudoinversion of various image degradation systems to attempt the removal of such degradation in an iterative fashion up to a point just prior to reaching singularity. Such approaches have a very general appeal as they can operate on both spatially invariant point spread function degradations as well as spatially variant point spread function degradations. Coupled with past image restoration activity is a current attempt to develop a posteriori approaches often referred to as blind deconvolution methods. These approaches measure the degradation of an image from the image itself with very limited or small amounts of apriori knowledge. While such approaches have had some fruition in the recent

past, it is expected that greater insight will be developed from ongoing current research. Again, time does not permit going into greater details in the image restoration projects but subsequent discussions will fill in the gaps.

The third area of interest at the Image Processing Institute at USC involves the application of mathematical problems for image analysis. Two particular topics that come to mind that are of considerable current interest involve the areas of degrees of freedom of imaging systems and finite field mathematics for signal processing of images. The former topic is one that is devoted to a mathematical analysis in an attempt to optimize the amount of information when obtained from various imaging systems and imaging configurations. By this it is meant that if one designs an imaging system to optimally take advantage of the degrees of freedom available from such a system, one will soon come to realize that the vast amounts of data that many imaging systems gather today are in fact tremendously redundant. By the use of Gramian techniques and degree of freedom analysis it is possible to hypothesize far superior data gathering approaches for a fixed number of data samples than are currently in existence. Examples of successes in this area involve analysis of three dimensional tomographic scanners as well as various NASA related ERTS satellite imaging systems. The future looks particularly bright in this area of analysis of borrowing from the field of mathematics and the variable knot approaches that that discipline supplies. This allows us to envision the possibility of variable Nyquist or Shannon sampling theorems resulting in two dimensional optimal scanning systems matched to the imaging systems in such a way that every data sample carries more information than might

traditionally be expected. The other area of interest in the mathematical problems for image analysis refers to the finite field mathematics of signal processing that has just recently come to the forefront in the area of finite field fast Fourier transform technology. Finite field arithmetic allows for perfect signal processing results without the problems of computational noise due to floating point and round off errors. The possibility of taking perfect Fourier transforms and perfect signal processing results by finite field algebra opens up whole new horizons of signal processing capabilities by now allowing for numerical instabilities to be further reduced without the need for more complicated or more sophisticated computer hardware. We're just scratching the surface in this area but it does look to be a profitable one for the future.

In closing I would like to simply say that in my humble opinion the problems of image processing of the future are much more oriented towards the exploitation of data rather than the sensing and gathering of data. We are all too familar with the many warehouses full of magnetic tapes of imagery that have not been analyzed because man has unfortunately not been able to build exploitation facilities as successfully as he has been able to build sensing equipment. I would submit that image processing successes of the future will be those in which very simple and very efficient techniques are developed for exploitation of imagery prior to transmission of imagery, for as we all know, no matter what the bandwidth of one's channel, that will always be the limiting constraint because of man's inherent belief "that if a little data is good, a lot of data must be better".

SOME DATA PROCESSING PROBLEMS RELATED TO THE ESTABLISHMENT OF A MULTI-SENSOR DATA BASE

Walter E. Boge
US Army Engineer Topographic Laboratories, Ft. Belvoir, VA

Samuel S. Rifman
TRW Systems and Energy, Redondo Beach, CA

ABSTRACT

Within the last decade or so there has been seen an increasing proliferation of different sensors and different sensor platforms designed for earth resources applications, ranging from aircraft borne sensors (including multichannel spectrometers, numerous RB-57 and U-2 imagery and even synthetic aperture radar) to satellite borne sensors (the LANDSAT series, Skylab, Apollo, etc.). The result of this collection effort is a great body of data, most of it residing in readily accessible facilities such as the EROS Data Center, Sioux Falls, South Dakota. Coincident with the creation of this treasure house of data has been the emergence of digital processing techniques using sophisticated mini computers, special processors, display systems and film hard copy devices. But has the digital technology for processing and digesting remotely sensed data kept pace with the collection process? Can we in fact efficiently process the enormous multi-sensor data base generated by satellite and aircraft collection systems for a wide range of applications, such as earth resources?

It is the purpose of this abstract to identify a few of the data processing problems related to the establishment of a multi-sensor data base. Recognizing that questions of national policy, coordination and cooperation involving numerous national programs may impact the nature and scope of the data collection process as well as the processing task (Ref. 1), it is nonetheless beyond the scope of the present discussion to treat such issues here. Rather, we shall assume an extant multi-sensor collection process, and shall address the problem of processing the resultant data base for a wide-range of applications.

Perhaps the first question to be addressed is how to merge the collection processes and the data requirements associated with particular functional applications. In other words, how can we best determine the set or sets of collection parameters (resolution, scale, spectral bands, band widths, periodicity, etc.) that characterize the multispectral data base which will best permit an analyst to extract the desired information? To accomplish this it will be necessary to establish specific data base requirements associated with specific application areas and determine how to amalgamate these requirements with minimal conflict. For example, the Army Corp of Engineers is charged with overseeing and maintaining all inland navigable waterways in the United States. If the monitoring function associated with this responsibility is ultimately to be accomplished using remotely sensed data from say a LANDSAT collection system; then, what are the unique data base requirements necessary to accomplish this mission? Also, how do these requirements relate to the requirements of other potential users of this type data, such as civil enterprises (e.g. mining, geology, etc.) and other governmental activities (e.g. forest inventory, polution control, etc.). If users needs can be defined and catalogued, there will then be a better basis for collecting, storing and using data of value and throwing away that which has no contribution. Unless we can set up efficient and effective collection and purging processes, we will sooner or later be buried with digits.

The next question is rather obvious: assuming all the data collected is useable and useful, how do we prevent the data processing activity itself from becoming a bottleneck? For example, only 30 minutes operation of just one of the 100 Mbps sensors projected for LANDSAT-D will double the 10^{11} bit/day input capacity of the Master Data Processor (MDP) for image geometric correction, which NASA expects to have operational by the end of 1977. To handle all the other sensor systems, must we replicate MDPs, or can we find alternatives? Is data compression the answer (currently lossless compression factors $\approx$ 2:1 are possible)? What new hardware technologies can be utilized for mass memories, magnetic tape drives and processors? In particular, can we develop relatively low cost processors and peripherals so that individual users can exploit with in-house capabilities the volume of data generated?

Having a multi-sensor data base immediately suggests the requirement to reference all the data to a common coordinate system. If this were done, the possibilities for multi-spectral and multi-temporal analyses would be greatly augmented. On the other hand, what common "universal" coordinate system should be employed?

Recently we have seen in one area the use of a map projection peculiar (or "natural") to the data collection system, viz., the Space Oblique Mercator projection for LANDSAT imagery (Ref. 2). The problem of defining a grid of image pixels over the surface of the earth acceptable to a wide class of users of remotely collected data (from a wide variety of sources) is certainly formidable. Equally formidable is the mathematical difficulty of making such a grid "uniformly dense" over the surface of the earth.

Assuming a universal map projection can be defined, is there available sufficient auxiliary data (related ultimately to sensor position and orientation) to make possible reasonable scene to ground and scene to scene spatial registration accuracies? Some early results achieved using LANDSAT data have demonstrated that under reasonable conditions the case of low resolution satellite borne sensors is readily amenable to solution (Ref. 3 and Ref. 4). In the case of high resolution aircraft borne sensors, this registration problem becomes much more difficult due to the platform dynamics and relief displacement. Thus several more questions arise: Do we need to instrument sensor systems more completely (and at greater cost), must we build more elaborate attitude-control systems, or will it be necessary to put greater effort into the ground processing system (possibly at the expense of processing speed)?

REFERENCES

1. See, for example, the first six articles in February, 1976 issue of Photogrammetric Engineering and Remote Sensing, Vol. 42, No. 2.

2. Colvocoresses, A.P., "Map Projection of the Bulk (System Corrected) ERTS MSS Image," Memorandum for the Record (EC-18-ERTS), Department of the Interior, Geological Survey (August 1, 1973).

3. Caron, R. H., "Evaluation of Full-Scene Registered ERTS MSS Imagery Using a Multitemporal/Multispectral Bayes Supervised Classifier," Proceedings of Fourth Annual Remote Sensing of Earth Resources Conference, University of Tennessee Space Institute, Tullahoma (March 24-26, 1975).

4. Rifman, S. S., et al, "Second Generation Digital Techniques for Processing LANDSAT MSS Data, "Proceedings of Earth Resources Survey Symposium, Shamrock Hilton Hotel, Houston, Texas, (June 8-13, 1975).

MAN-MACHINE INTERFACE IN IMAGE PROCESSING

Andrew G. Tescher

The Aerospace Corporation
P.O. Box 92957
Los Angeles, CA   90009

ABSTRACT

Computer technology has had a strong impact on image processing. However, the advances in image processing technology have been somewhat restricted both in scope and user community. Perhaps the most useful area of numerical image processing has been the correction of geometric and photometric image defects introduced by the limited quality image sensor. The impressive results obtained for this type of application, for example at JPL, has been widely acknowledged. On the other hand, the analytical methods have been image independent, and limited user interaction was required. Furthermore, accurate model of the distortion parameters was also required.

I would like to address two additional areas of image processing. For lack of better expression, "practical" pictorial pattern recognition will be considered as relevant to various applications, such as image coding and biomedical image processing. The second proposed application is the role of parameterization in image processing. This general technique includes various procedures by which the "untrained" analyst can interact with the image processing software. Here, the fundamental difficulty is how to represent relatively complicated image processing functions by few parameters. Several potential solutions, particularly in the area of filtering, will be discussed.

EROS DATA CENTER IMAGE PROCESSING TECHNOLOGY

Allen H. Watkins

EROS Data Center
U. S. Geological Survey
Department of the Interior
Sioux Falls, South Dakota

ABSTRACT

Over the past 18 months, the EROS (Earth Resources Observation Systems) Data Center of the U.S. Geological Survey, working with the Jet Propulsion Laboratory and the Geological Survey's Center for Astrogeology in Flagstaff, New Mexico, has adapted algorithms for digital image enhancement and processing and is currently installing these algorithms in production processing equipment at the Center in Sioux Falls, South Dakota. Products resulting from these processing systems will be available to the public in very limited quantities early in 1978 and in production quantities for Landsat-C data. Landsat-C is scheduled to be launched by NASA in late 1977. This paper describes many of the techniques being employed, along with a description of current and planned future systems, and presents examples of resulting products, comparing them with existing image products from Landsat-1 and -2.

EROS DATA CENTER IMAGE PROCESSING TECHNOLOGY

The Earth Resources Observation Systems (EROS) Data Center, is administered by the U.S. Geological Survey for the United States Department of the Interior. The Center is a principal operating element of the EROS Program, which was created in 1966 to apply remote sensing techniques to resources inventory, monitoring, and management. The Center is located 16 miles northeast of Sioux Falls, South Dakota, in a modern 125,000 square foot facility. The Center provides products in the form of remote sensing data acquired from NASA satellite and aircraft projects, along with aerial mapping photography held by the Department of the Interior. The Center also provides assistance in the application of remotely sensed data to resource inventory, monitoring, and management activities. This assistance takes the form of both domestic and international training, the conduct and documentation of cooperative applications projects, and day-to-day applications assistance in the use of special analysis equipment.

The EROS Data Center was established in 1972, with the move into the new Center complex in early 1974. Sioux Falls, South Dakota, was selected as the location for the new facility because of its central location and potential to receive telemetry data from satellites in a Landsat-type orbit in real time while over any portion of the lower 48 states.

Data holdings at the Center include over 800,000 frames of Landsat imagery; Landsat electronic data in the form of computer compatible tapes; over 50,000 frames of Skylab, Apollo, and Gemini spacecraft data; more than 1,500,000 frames of data from the NASA research aircraft program; and almost 3,000,000 frames of Department of the Interior aerial mapping photography. Products

from this data base are available in color or black and white photographs, in sizes from 16 millimeter through 40 inches, and computer compatible tapes in 7 or 9 track, and 800 or 1600 bit per inch formats. In addition, data catalogs and specialized computer search listings are available upon request.

At the heart of the Data Center is a central computer complex, which controls and accesses a data base of the more than 6,000,000 frames of Earth resources data, performs searches of specific geographic areas of interest, and serves as a management tool for the entire data reproduction process. The computerized data storage and retrieval system is based on a geographic system of latitude and longitude, supplemented by such information as image quality, cloud cover, and type of data. A customer's inquiry as to data availability may be a geographic point location or a rectangular area bounded by latitude and longitude. Based on the customer requirements and the use to be made of the data, a computer geographic search will print out a listing of available imagery from which a final selection can be made.

The demand for remotely sensed data from the EROS Data Center has continued to increase both in volume of frames and dollar value. Table 1 indicates the growth in demand from 165,000 frames of data in FY 73 to almost 500,000 frames in FY 76 and dollar sales volume growth from $375,000 in FY 73 to almost $2,600,000 in FY 76. The fastest growing demand is for Landsat electronic data in the form of computer compatible tapes.

Table 2 is a statistical summary of customer types purchasing remotely sensed data from the Center and, as can be seen, U.S. industry and agencies of the U.S. Government are the largest users.

A detailed look at the principal users and purchases of EDC products reveals that almost all categories of U.S. industry and a wide variety of Federal, State, and local government agencies are currently using the data for a variety of applications. The principal application of the data appears to be mineral and fossil fuel exploration and related geologic base mapping activities of the major petroleum and mining companies. This area would appear to account for almost 50 percent of total sales of data from the Center.

The average delivery time for data is approximately 2 weeks following receipt of the order at the Data Center, with a median delivery time of 10 days. During 1975, 76 percent, or 16,405 orders, were shipped prior to 14.7 calendar days. More than 20,000 of a total 21,500 orders were shipped in less than 33 days.

All products from the Center are inspected by a quality control staff prior to shipment. Customer feedback in the form of prepaid postcards included with each shipment have consistently indicated good or better quality for approximately 90 percent of orders shipped.

Inquiries concerning availability of data in the EROS Data Center result in a geographic computer search to determine data available over a specified area of interest. This computer search can be requested by a letter, visit, or phone call to either the EROS Data Center or one of the EROS Applications Assistance Facilities located throughout the country. An inquiry form is available and may be used to request information on available coverage over a specific area of interest. Requests for a geographic search may take the form of a point search where all images over any portion of the point will be included or a rectangular method where any area bounded by latitude and longitude can be described. An eight-sided polygon can also be used to specify the geographic area of interest. Latitude and longitude coordinate specification is preferred since this is the method required for the computer geographic search. Additional information must be included covering acceptable dates and seasons of exposure, type of imagery preferred, cloud cover, quality, etc. Geographic areas must be clearly identified and should be limited in size as much as possible to avoid potentially voluminous output and the need to interpret large numbers of choices.

Orders for reproductions of data from the EROS Data Center are accepted from individuals, government organizations, universities, and industries from the United States and all foreign countries. All orders must be accompanied by check, money order, purchase order, or authorized account identification, and processing cannot be initiated until correct payment is received.

Standing or open accounts may be established by repetitive users. To open a standing account, a check must be remitted for the amount to be deposited. Customers will be informed of the account number and future orders can be placed referencing this open account number.

Custom processing to unique scales, formats, and processing criteria is available from the Data Center. These services normally require longer periods of time for completion and pricing is based on three times the standard price. In addition, a priority system for rapid delivery of standard products is available whereby products will be shipped within 5 days from receipt of order. To provide this service, a three times standard price is charged and priority processing will only be accepted where imagery is specifically identified, standard products ordered, and payment is enclosed or credit available.

In addition to the central facilities near Sioux Falls, South Dakota, the EROS Data Center operates a number of EROS Applications Assistance Facilities where the public may view microfilm copies of imagery available from the Data Center and can receive assistance in the research, ordering, and techniques of applying the data to resource problems. These Applications Assistance Facilities maintain microfilm copies and provide computer terminal inquiry and order capability to the central computer complex at the EROS Data Center. Facilities are currently established in Menlo Park, California; Phoenix, Arizona; Reston, Virginia; Bay St. Louis, Mississippi; Ft. Clayton, Canal Zone; and Fairbanks, Alaska. In addition to these offices, small Data Reference Files have been established throughout the United States to maintain microfilm copies of the most used data available from the Center and to provide assistance to the visitor in reviewing and ordering data. The Data Center also functions as an integral part of the National Cartographic Information Center (NCIC) which provides integrated mapping, photographic, and geodetic control information and data. A number of NCIC offices are connected to the central computer complex via remote terminal throughout the country.

Training in remote sensing and assistance in techniques for extracting information for remote sensing data are available at the EROS Data Center. Inquiries can be made by telephone, letter, or personal visit to the Applications Branch at the Center.

Periodic training programs in remote sensing are given at the Center. Normally, these programs are several days to 1 week long and stress the use of data in a particular application, for example, agricultural inventory, water management, etc. Two or three times a year, a 3 to 4 week course is offered to foreign nationals, stressing the fundamentals of remote sensing with exposure to a number of applications areas. The principal role of the EROS Data Center in applications assistance is technology transfer.

The Data Analysis Laboratory is another important part of the EROS Data Center training and assistance program. The Laboratory staff is responsible for developing, maintaining and operating state-of-the-art automatic and semi-automatic digital and analog image analysis hardware/software systems in support of all technology transfer programs. The Laboratory is a unique data analysis facility within which Earth scientists from the EROS Data Center, EROS Program Office, USGS and other user organizations can receive training, and demonstrations, and perform evaluations using digital maximum likelihood classification algorithms, thematic extraction using parallelepiped functions, and video disk storage and retrieval techniques.

The following image analysis hardware/software systems have been incorporated into the Laboratory to date:

GE Image 100 Analysis System: The system is an interactive multispectral digital image analysis system consisting of two (2) tape drives, PDP 11/35 minicomputer, INTEL solid state memory, interactive terminal, user console, color CRT display and line printer. Over one-hundred software functions are available to electronically process, enhance, classify, analyze and display digital image data.

ISI-170 Digital/Analog Image Analysis System: When fully developed, this system will provide the capability of storing up to 256 multispectral, multidate images for real time analysis. The system consists of a vidicon TV scanner, tape drive, PDP 11/20 minicomputer, video disc, analog encoder,

density slicer, color and isometric CRT displays. The system incorporates the advantages of analog and digital image analysis techniques.

ESL IDEMS System: This digital image analysis system provides the capability of large area analysis and classification using maximum likelihood and clustering algorithms. The system components include an image analysis preprocessor for training set statistics compilation and the EROS Data Center general purpose computer for large area analysis. The preprocessor consists of two tape drives, HP 3000 minicomputer, Comtal image processing display, two interactive terminals and a high speed line printer. After analysis of training sets within image subscenes has been performed on the preprocessor, the digital image data tape is analyzed on the Burroughs B6700 computer using the statistics generated on the training sets. The LARSYS maximum likelihood classification system provides higher quantitative accuracies than smaller interactive image analysis systems.

The Laboratory serves as a national/international focal point for leading Earth scientists and representatives of user organizations to interact with data analysis specialists on applications of current state-of-the-art techniques to a variety of Earth resource problems.

Within the next few months, the EROS Data Center will start to make available enhanced Landsat image and CCT products in limited quantities. Prices for these user unique products will be established to cover the cost of processing and production and may well be on the order of $750 to $1000 per Landsat scene.

The following corrections and enhancements will be available for both images and CCTs:

Earth Rotation Correction (deskew). The last set of 6 scan lines of the scene is treated as the base and the first 2250 lines are translated relative to it. Each line is translated to the right by a number of pixels which is equal to the integer part of TRAN (LINE) computer as follows:
(1) TRAN (LINE) = (NUMLIN-LINE) (MAXT/NUMLIN-1)
   Where: NUMLIN is the number of lines in the scene (2256) divided by 6
   LINE is the number of the line being processed divided by 6
   MAXT is defined to be $k°$ Cos (H) $°$ Cos (Lat) where K = 230.564944
      Cos (H) is cosine of spacecraft heading minus 180°
      Cos (Lat) is the cosine of the center latitude
(2) The constant 230.564944 has been determined by empirical methods to be the constant needed to provide the proper translation at the equator.

The effect of this process is to take a scene which is square on a CCT and image it on film in the shape of a parallelogram such that terrain features are in proper position relative to each other from top to bottom of the scene.

Line Length Adjustment. The length of each line is adjusted to a length of 3318 pixels. This is done by adding additional pixels to the line after GSFC has added pixels to bring the length to 3240. Mirror speed correction is also made during the

line length adjustment.
(1) That is OLOC = (ILOC) $\frac{ADJLEN}{ORGLEN}$
   where OLOC = output pixel location
      ILOC = input pixel location (unadjusted)
      ADJLEN = adjusted line length
      ORGLEN = unadjusted line length

To determine the value of the pixel a weighted ratio of the distance between the pixels is computed. OVAL = $\chi_1 \lambda + \gamma_2 (1-\lambda)$

(2) Each line is corrected for length by first dividing it into 8 segments. GSFC has provided constants which are used in determining mirror speed correction for each of these 8 segments. These constants determine which segments need to have pixels added. The line length adjustment is further made by adding pixels to each segment to bring the total to 3318. This is accomplished by calculating pixel locations in the adjusted line, then comparing the original line to it. Pixels are then added to fill gaps where they occur.

Destriping.

(1) The first step is to modify the input pixel as follows: $O = Ma + \frac{Sa}{Si} . (I - mi)$
   where O = output brightness value
      Ma = mean of the entire scene
      Sa = standard deviation of the entire scene
      Si = standard deviation of the current scanner only
      I = input brightness value
      mi = mean of the current scanner only

Then, local averages are compared:

$$DIFF = LOCAV_i - LOCAV_{i-1}$$

   where DIFF is the difference in local average
      $LOCAV_i$ is the local average of the current line about the pixel being processed
      $LOCAV_{i-1}$ is the local average of the previous line about the corresponding pixel

   if $|DIFF| < MODLIM$
   then $Q(i) = P(i) - DIFF$
   where Q(i) is the processed pixel brightness value
      P(i) is the unprocessed pixel brightness value
   if $|DIFF| \geq MODLIM$, a naturally occuring edge is assumed to be present and no change in brightness value is made.

Also, before these calculations are made, the unprocessed values (current pixel with the corresponding pixel in the previous line) are compared. If the difference is greater than 30, a diagonal edge is assumed to occur in the local vicinity and these values are not used in computing average values for the filter.

(2) Line numbers 43, 50, 57, 64, etc., are not destriped. However, local averages are computed at each pixel in these lines to be used in the destriping process. A line is destriped by computing at each pixel a local average of NFIL

(an odd integer) points (presently NFIL = 151) about the pixel being processed. The local average is compared with the corresponding local average of the next lower numbered line. If the magnitude of the difference in the local averages is less than a modification limit (MODLIM, 8 is the currently used value), then the pixel brightness value is modified by the difference.

Contrast Stretch. Contrast stretching is performed by computing for each unprocessed brightness value a new value obtained as follows:

(1) $O(i) = \dfrac{I(i)-INMIN}{INMAX-INMIN} * 255$

> Where $O(i)$ = the output brightness value of the ith pixel
>
> $I(i)$ = the input (unprocessed) brightness value of the ith pixel
>
> INMIN = the minimum pixel brightness value in the scene (or a minimum user input value)
>
> INMAX = the maximum pixel brightness value in the scene (or a maximum user input value)

(2) This contrast stretch is a linear mapping which maps brightness values equal to INMIN to zero and brightness values equal to INMAX to 255. All values between INMIN and INMAX have a constant multiplier applied to them which linearly spreads out the pixel values.

Enhance Filter (Edge Enhancement). The processed brightness value, $Q(i)$, of the ith sample of the line being processed is generated from the unprocessed brightness value, $P(i)$ as follows:

(1) A local average, $A(i)$, about location i is computed by averaging N points on the line surrounding i

> that is $A(i) = \dfrac{1}{N} \sum\limits_{k=i-N2}^{i+N2} P(k)$

> where $A(i)$ = the average about the ith pixel
>
> N = the number of pixels to average
>
> N2 = the integer part of N/2
>
> $P(k)$ = unprocessed brightness value of the kth pixel

> then $Q(i)$ is computed as a weighted sum of $P(i)$ and $A(i)$
>
> $Q(i) = P(i) + [P(i) - C°A(i)]$

(2) That is, the processed brightness value is equal to the unprocessed value plus an enhancement factor. This enhancement factor is the difference of the unprocessed value and the product of a constant and the local average. (The present form of the program automatically uses N equal to 11 and C equal to 1). The effect is to make those pixels which are brighter than the local average brighter yet and those which are darker, darker.

For Landsat-3, now scheduled for launch in late 1977, NASA plans to implement an all digital processing system and provide the EDC High Density Digital Tape data. The EROS Data Center in turn will implement a Production Digital Image Processing System that will provide corrected and enhanced digital and photographic products to users in a more timely and efficient fashion. The EROS Data Center system will consist of three major subsystems as follows: 1) Incoming Inspection Subsystem (IISS) - that will screen for data quality and cloud cover, produce browse and quicklook film, generate an edited master tape and generate information for the main image file data base; 2) Production Processing Subsystem (PPSS) - that will perform geometric and radiometric corrections, transform data to specific map projections, and generate a corrected tape for product generation; 3) Product Generation Subsystem (PGSS) - that will convert tape to film, reproduce and reformat both high density and low density digital tape data, and perform specific scene enhancements.

In addition, NASA and the Department of the Interior are studying concepts for the use of communication satellites for data relay to significantly improve the timeliness of data dissemination. This concept includes the reception of data at the NASA Alaskan, Goldstone, and Goddard reception stations, with subsequent relay to Sioux Falls on a near-real time basis for further processing, reproduction, and dissemination to users.

To obtain further information contact the EROS Data Center, (605) 594-6511.

Table 1.   EDC DATA DEMAND
July 19, 1976

	FY 73 Items	FY 74 Items	FY 74 % Increase	FY 75 Items	FY 75 % Increase	FY 76 Items	FY 76 % Increase
I. FRAMES							
Landsat Imagery	81,071	157,178	+ 94%	195,125	+ 24%	246,449	+ 26%
Landsat CCT's	10	228	+ 2,000%	879	+ 286%	2,305	+ 162%
Manned Satellite Imagery	--	17,201	--	28,049	+ 63%	9,664	- 66%
Aircraft Imagery	83,942	109,490	+ 30%	190,031	+ 74%	148,993	- 22%
TOTALS	165,023	284,097	+ 72%	414,084	+ 46%	407,411	- 2%

	FY 73 Dollars	FY 74 Dollars	FY 74 % Increase	FY 75 Dollars	FY 75 % Increase	FY 76 Dollars	FY 76 % Increase
II. DOLLARS							
Landsat Imagery	$228,042	$528,514	+ 132%	$ 760,263	+ 44%	$1,237,862	+ 63%
Landsat CCT's	1,600	36,480	+ 2,200%	169,300	+ 364%	405,942	+ 140%
Manned Satellite Imagery	--	34,421	--	113,473	+ 230%	85,783	- 24%
Aircraft Imagery	144,676	237,332	+ 64%	566,806	+ 139%	734,967	+ 30%
TOTALS	$374,318	$836,747	+ 124%	$1,609,842	+ 92%	$2,464,554	+ 53%

Table 2.   CUSTOMER PROFILE - FY 76*

	% Frames	% Dollars
Landsat Imagery		
Federal Government	17	20
State/Local Government	1	1
Academic	14	12
Industrial	23	22
Individuals	10	14
Non-U.S.	35	31
Landsat CCT's		
Federal Government	32	19
State/Local Government	0	0
Adademic	15	18
Industrial	29	34
Individuals	2	3
Non-U.S.	22	26
Other Imagery		
Federal Government	41	36
State/Local Government	9	7
Academic	9	10
Industrial	27	27
Individuals	12	16
Non-U.S.	2	4

*NASA Principal Investigators Excluded.
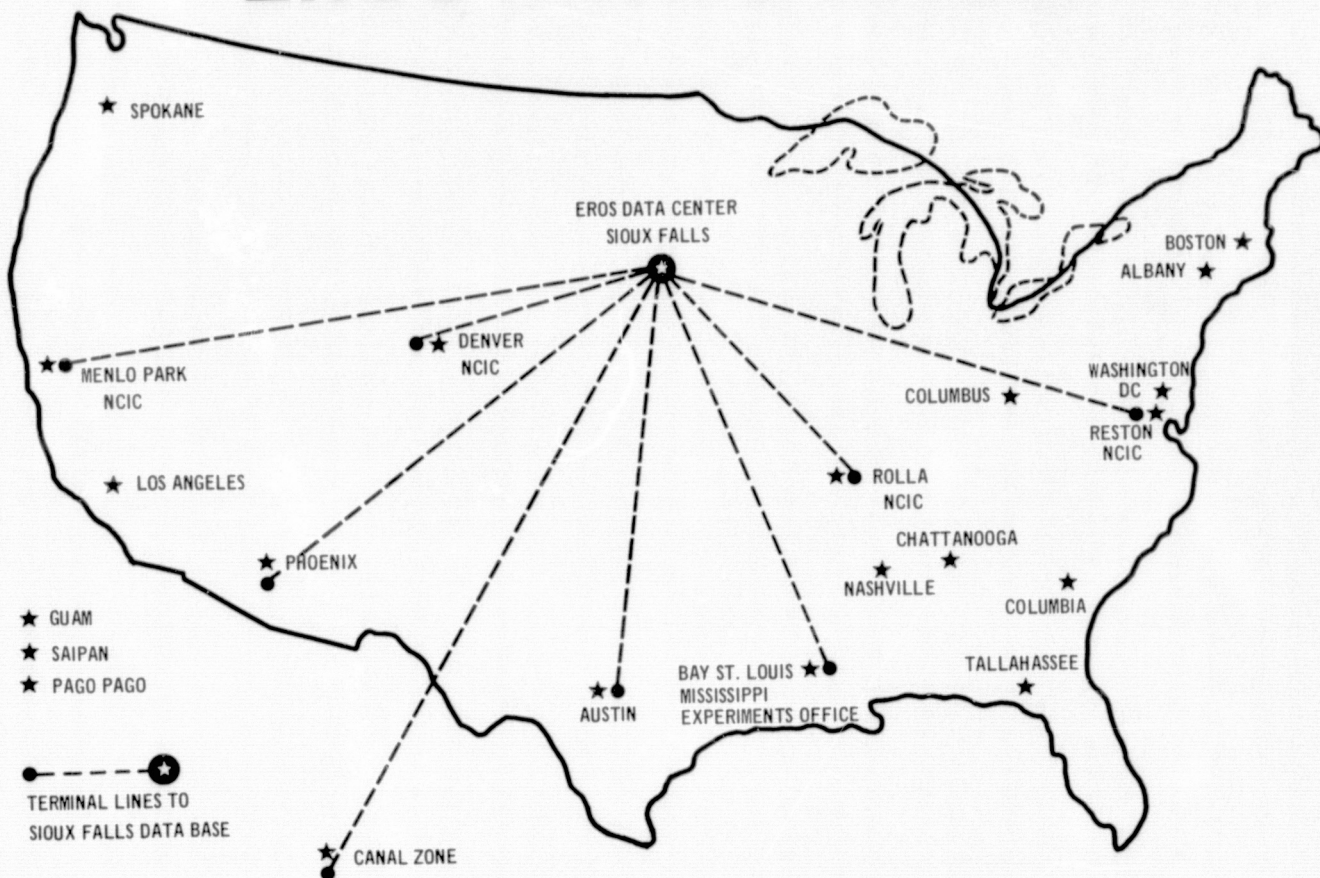
EROS DATA SYSTEM

Fig. 1 EROS Data Center Network

BROWSE FILM
QUICK LOOK FILM

CCT

HDT$_A$
(FROM
NASA)

| INCOMING INSPECTION SUBSYSTEM (IISS) | | PRODUCTION PROCESSING SUBSYSTEM (PPSS) | | PRODUCT GENERATION SUBSYSTEM (PGSS) |

HDT$_{EM}$

HDT$_P$

HDT$_P$

CCT$_P$

FILM PRODUCTS

- Screen for Data Quality and Cloud Cover
- Produce Browse and Quick Look Film
- Generate Edited Master (HDT$_{EM}$)
- Generate Information for Main Image File Data Base

- Perform Geometric Corrections
- Perform Radiometric Corrections
- Apply Map Projection
- Generate HDT$_P$

- Convert Tape to Film
- Convert HDT's to Different Formats
- Duplicate HDTs
- Convert HDT's to CCT's
- Convert CCT's to HDT's
- Perform Scene Enhancements

Fig. 2. EROS Data Center Digital Image Processing System (EDIPS)

- Geometric Corrections

    - Earth Rotation/Skew
    - Aspect Ratio
    - Mirror Scan Non Linearity

- Noise Removal

    - Destriping
    - Dropout Smoothing

- Scene Enhancements

    - Contrast Stretch
    - Radiance Level Shift
    - Haze Removal
    - Edge Sharpening
    - Sun Angle Correction
    - 1:1,000,000 Film Original

Fig. 3. Digital Image Corrections and Enhancements

IMAGE PROCESSING SOFTWARE AVAILABLE THROUGH COSMIC

Walter B. McRae

Computer Software Management and Information Center (COSMIC)
Athens, Georgia

ABSTRACT

The Computer Software Management and Information Center. COSMIC, is a computer program collection, evaluation, and distribution facility operated by the University of Georgia under contract to the National Aeronautics and Space Administration in support of the NASA Technology Utilization Program. As a consequence of NASA efforts to increase utilization of LANDSAT collected earth resources data, a number of computer programs for use in the correction, analysis, and display of digitized image data are now available through COSMIC. A brief description of some of these NASA sponsored digital image processing systems is presented in this paper.

I. INTRODUCTION

The Computer Software Management and Information Center, COSMIC, is a software collection and distribution center operated by the University of Georgia under contract to the National Aeronautics and Space Administration. Established in 1966, COSMIC is the dissemination center within the NASA Technology Utilization program responsible for making available to the American public computer programs developed in NASA sponsored research and development efforts. Although early activities were restricted to the distribution of NASA developed computer programs exclusively, several other federal agencies now also utilize COSMIC as the distribution center for software developed under their auspices. Some of these other agencies are the Department of Defense, the Veterans Administration, and the Library of Congress. A limited number of computer programs acquired from foreign sources in exchange for NASA software are also distributed by COSMIC.

Since inception in 1966, over 3,500 programs have been supplied to COSMIC for distribution. Of this number, approximately 1,500 are currently available to the public with the remainder being rejected, typically for reasons of incompleteness, in the verification and evaluation functions COSMIC performs.

The COSMIC facility by contract requirement must operate on a cost recovery basis. Consequently, fees are charged for the distributed materials in order to support a staff of approximately 20 technical and support personnel with prices established according to a NASA authorized pricing schedule. The average cost of program source code available through COSMIC is $365 with the mean cost of supporting documentation, which is available separately from the program code, being $19. Most distributions are by sale, however, a growing number of distributions are through lease agreements administered by COSMIC. In particular, two of the digital image data processing systems discussed below are available on a lease basis only.

Computer programs available through COSMIC are announced in the quarterly journal 'Computer Program Abstracts' published by NASA which is available on a subscription basis from the Government Printing Office or directly from COSMIC for an annual subscription fee of $10 to domestic subscribers and $15 to foreign customers.

Software distributed by COSMIC typically deals with applications in the engineering, mathematical, and computer sciences because of the originating sources for the programs. The variety of available programs is increasing, however, as NASA research activities broaden. This diversification has been especially evident in the geosciences in recent years with the acquisition of a number of programs for processing LANDSAT recorded digital earth resources data. The remainder of this paper is devoted to summary descriptions of several of the most popular digital image processing systmes currently distributed by COSMIC.

II. PROGRAM DESCRIPTIONS

1. LARSYS-Multispectral Data Analysis System (Release 3.1)

LARSYS was developed at Purdue University under a research project sponsored by the U. S. Department of Agriculture. The primary objective of this research was to develop remote sensing techniques and investigate their application to the field of agriculture; however, applicability has also been demonstrated in the area of geology, hydrology, and geography. The primary input data to LARSYS is multispectral data in image orientation obtained from aircraft mounted multispectral scanners; digitized photography from spacecraft overlayed to create registered multispectral images; multispectral data from the ERTS satellite; and other aircraft scanners.

The basic analysis concept of LARSYS consists of locating data points which are believed to be representative of classes of interest. A class of interest may be certain crops, beaches, woods, or

geological features. Gaussian statistics of these data points (a key assumption made in several LARSYS algorithms is that distributions are Gaussian) are calculated and data sets are classified by spectral similarity. Next, the classification results are evaluated. Thus, there are four basic concepts to the analysis

- location of data points
- statistical calculations
- classification
- process evaluation

The data points selected to calculate statistics are called training data points or training fields. The data points used for evaluation are called test fields.

The LARSYS software consists of eighteen processing functions written primarily in FORTRAN IV but with some coded in IBM Assembler language. These processes may be divided into the following six groupings.

- Utility functions that enable the user to extract or manipulate data from multispectral image storage tapes.
- Evaluation functions to assist the user in graphically displaying image data from selected scanlines, columns, and channels.
- Field Selection Functions to aid in the selection of training fields and test fields.
- Statistical Calculation Functions that implement a clustering algorithm to classify individual data points into a predefined number of clusters, provide statistics on training fields, and calculate the degree of separability between classes.
- Classification Functions used to classify data by implementing statistical distance measures and a maximum likelihood schema.
- Analysis Functions that enable the user to manipulate and analyze produced by the classification functions.

LARSYS is written for implementation on IBM 360 systems for use in either interactive or batch mode. The system was developed on an IBM 360/67 operating under the Cambride virtual machine monitoring system.

2. SMIPS-Small Interactive Image Processing System

SMIPS is designed to facilitate the acquisition, digital processing, and recording of image data as well as pattern recognition in an interactive mode. The system contains approximately 104 routines and application programs and is designed for ease of communication with the computer by personnel who are not expert pro-

grammers, fast response to requests for information on pictures, and complete error recovery. SMIPS is intended as an experimental system to provide rapid familiarization with the characteristics of image data rather than as a production system. It can, however, be used for production work and its full compatability with the VICAR System allows access to numerous image processing programs in the VICAR library. Because of its modular design, new interactive capabilities can be easily included and SMIPS may be used as an experimental tool to gain further experience for the design of larger and more sophisticated interactive image processing systems.

The SMIPS system has been developed to give the user flexible and convenient control of a variety of image processing methods including:

- Fast display of parts of pictorial data on the screen of an IBM 2250 display device either numerically or as a character representation.
- Computation and display of histograms.
- Convenient specification of a variety of image processing tasks for restoration, enhancement and detection.
- Output of numerical results and pictures as graphs, maps and photographs.

The SMIPS program is designed for implementation on IBM 360/370 series equipment with the delivered source code written in FORTRAN IV and IBM Assembler. The program has been installed at the Goddard Space Flight Center on an IBM 360/91 operating under OS/MVT.

3. ASTEP-Algorithmic Simulation Test and Evaluation Program

This system represents a data analysis program used to examine statistical properties of multispectral scanner data. It serves as a tool to perform experiments to gain understanding of the problems associated with processing multispectral earth resources data and to test and evaluate processing algorithms. Examples of the analyses available with ASTEP are:

- quantitatively determining the variation in spectral signatures for recorded data
- determining if there are patterns in the signature variation, either spectrally or spatially
- determining the statistical homogeneity of typical ground truth sites
- determining if the statistical assumptions required for maximum likelihood processing of typical areas are satisfied
- evaluating the performance of

clustering techniques
- comparing the performance of clustering and maximum likelihood algorithms.

The current release of ASTEP can accept input data in LARSC1, LARSC2, ERTS, and Universal formats, and output processed image or data tapes in Universal format. The program consists of over a hundred subroutines written entirely in UNIVAC FORTRAN V, operates in either interactive or batch mode, and has been implemented on the UNIVAC 1100 EXEC 8 computer.

4. ELLTAB-Elliptical Table Lookup Algorithm for Pattern Recognition

This program is used to classify remotely sensed data into categories representing types of vegetation, geological features, or other earth surface aspects of interest by applying an advanced table look-up approach to pattern recognition. The essence of the table look-up approach is that at classification time a remotely sensed unit or pixel is assigned to a category by merely looking up its channel readings in a table as opposed to making the lengthy calculations required in a maximum likelihood computation. The table look-up approach makes it possible to process multispectral scanner data using minicomputer (i.e., low cost, small scale, general purpose digital computers). Also, processing time is reduced significantly compared with the conventional approach on the same computer. Although the ELLTAB table look-up method requires much less core than previous table look-up methods, it still needs more than the minimum amount required to implement the maximum likelihood decision algorithm. ELLTAB is limited to a maximum of four channels and the current version requires exactly four channels. Also, confidence levels are not available at classification time and consequently cannot be output to the classification tape.

The program has been developed and installed on a UNIVAC 1108 computer operating under the EXEC 8 monitor and is written entirely in UNIVAC FORTRAN V. The program may be operated in both batch and demand mode.

5. LACIE-Large Area Crop Inventory Experiment Image Processing System

The LACIE system is designed to analyze data from various imaging sensors including the ERTS/LANDSAT satellite (both MSS and RBV data), the MSC 24-channel multispectral scanner, the EREP S-192 experiment, and the Michigan 12-channel scanner. The system can be easily modified to process data from additional sensors.

The primary analysis software packages in the LACIE system are the Pattern Recognition applications, which perform multispectral analysis, and the Image Registra-

tion application. The Pattern Recognition application applies stochastic processes to digitized multispectral image data to classify unknown data into distinct materials. This application is based on the LARSYS program. The Image Registration application provides the user with two basic capabilities: first, it allows the user to register two images together; and second, it allows the user to register an image to a latitude-longitude (UTM) grid. In either type of registration, the final output from the registration application is an image. This iamge can then be combined with other images in LACIE by using an Image Composition application, and/or it can then be processed by any other application in LACIE.

LACIE uses the IMS Data Base/Data Communications (DB/DC) as its primary data base manager and five different IMS data bases to store information. LACIE operates in batch mode or interactive mode. The LACIE system can be implemented on an IBM 360 Series computer operating under OS MVT Release 21, VS1, VS2, and MVS with the appropriate release of the IMS DB/DC program product. The system requires 300K bytes of main storage, not including OS storage requirements. The additional machine requirements are 230K bytes of storage for IMS, 600K bytes for large capacity storage (LCS), five 2314 disk drives, seven 9 track tape drives, and one cluster of digital television equipment.

6. DAM-Detection and Location of Surface Water

The Detection and Mapping (DAM) computer package is a remote sensing program designed to detect and map surface water using ERTS-1 and 2 multispectral scanner (MSS) data from system corrected Computer Compatible Tapes (CCT's). The system contains a number of specialized programs to perform different functions required in the production of an accurately scaled classification map and may be executed in either batch or demand mode.

Detection and mapping procedures using DAM involves four phases:

- Data Acquisition to assemble the required ERTS imagery and data tapes, maps, meteorological data, and ERTS data catalogs.
- Control Network Establishment to establish a network of ground control points identifiable on both the ERTS data and corresponding maps to insure a computer generated map which is accurately scaled and registered for correlation with a conventional map.
- Classification and Mapping resulting in computer generated line printer output showing the location of each point classified as water at the scale specified, a listing of selected incremental

geographic coordinates corres-
ponding to latitude/longitude sym-
bols on the computer generated
output, and an optional listing
of pixels classified as water.
- Information Correlation and Inter-
pretation to mosaic the individual
line printer output strips to form
a map and to transfer the computer
derived information to conventional
maps for further analysis.

The program operates in batch and
demand mode and can be conversational; it
has been implemented on the UNIVAC 1100
running under EXEC 8. Minimum storage
requirements are 20K words of core storage
and 900K words of user disc/drum.

7. QUIKLOOK-LANDSAT Multispectral
Scanner Image Processing Program

This is an image enhancement program
containing 14 subroutines to approximate
geometric and radiometric corrections to
LANDSAT data. The program allows the
user to quickly produce a magnetic tape
in a format suitable for use by the
DICOMED film recorder containing the data
for a specified band of a full or partial
LANDSAT multispectral scanner image.
QUIKLOOK can also output the data on mag-
netic tape for further image processing
by the SMIPS program. Geometric correc-
tions performed by QUIKLOOK include cal-
culation of the earth rotation (skew)
correction, mirror velocity profile error,
and aspect ratio distortions. QUIKLOOK
also corrects radiometric values for non-
normal sun elevation angle.

The program only performs approximate
geometric and radiometric correction and
its output should not be construed as
being precision processed; the relative
geometric error in a QUIKLOOK output image
is greater than 300 meters. The methods
of correction used are those which can be
done computationally in a very short time.
The program operates in batch mode,
requires 128K bytes of storage, and has
been implemented on the IBM 360/91.

III. CONCLUSION

The digital image data processing sys-
tems described in Section II are each
currently available from COSMIC. For
further information on their capabilities
as well as procurement procedures and
costs, contact COSMIC at (404) 542-3265
or write: COSMIC, Suite 112, Barrow Hall,
Athens, Georgia, 30602.

PANEL ABSTRACTS

"ACTUAL USER EXPERIENCE AND NEEDS"

Michel T. Halbouty, Chairman

Participants:

Miner B. Long

Leo J. Miller

Jon W. Davidson

William W. V. Trollinger

Keith Fullenwider, Sr.

ACTUAL USER EXPERIENCE AND NEEDS

Michel T. Halbouty

Consulting Geologist
and
Petroleum Engineer
Houston, Texas

ABSTRACT OF OPENING REMARKS

I have previously made two statements which I repeat here. "The Landsat program is the most significant project ever flown by NASA", and "It is one of the most important national initiatives ever conceived and implemented in our 200 years of experience.

The project has enabled space technology to be merged with advances in earth science concepts, and this is occurring at a time when such an association is most needed to assist in the search for new energy and mineral resources.

Of course, technology for earth satellites will continually be improved and advanced; however, what we are interested in in this session is to discuss the application and implementation of the Landsat data to practical uses in the area of geology and geophysics of energy and mineral resources hydrology and water resources, oceanography and marine resources, geography, cartology and cultural resources and agricultural and forestry production. It is an established fact that the Landsat program has been of major importance to the earth scientist. Geological structures and land forms that had not been identified previously are now very obvious, and it is from these new identifications and interpretations of the surface of the earth which will lead to greater discoveries of our needed energy and mineral resources.

It should be stated clearly that satellite data are not the entire answer to problems of exploration or monitoring, but rather constitute a new element of information to be coupled with existing technology to increase the efficiency of search. This statement can certainly be illustrated by the fact that without remote sensing data, the cost of man-conducted ground survey to establish a resource target would be very high. If man were able to obtain aircraft remote sensing data, the cost of conducting an earth survey to establish a resource target would be much less. On the other hand, if man were able to obtain spacecraft remote sensing information, the overall cost of the ground survey to establish a resource target would be still less.

IMAGE PROCESSING TECHNOLOGY AND PETROLEUM EXPLORATION

Miner B. Long

Shell Oil Company
Houston, Texas

ABSTRACT

In the field of petroleum exploration the principal impact of additional image processing capability probably lies in improved methods of detection of surface geochemical anomalies. The demonstrated value of these anomalies in petroleum exploration, however, is not clear. Our ability to map geochemical anomalies using surface methods already far exceeds our ability to interpret them.

To be useful in petroleum exploration, a geochemical anomaly must result from the escape of hydrocarbons from a subsurface trap in which case their presence will lead us to undiscovered accumulations. The research we need lies in understanding the origin of these anomalies in order to determine if a direct relationship exists with subsurface accumulations.

THE NEED FOR ORE CONTROL IDENTIFICATION FROM IMAGE PROCESSING TECHNOLOGY

Leo J. Miller

Texasgulf, Inc.
Golden, Colorado

ABSTRACT

Historically, remote sensing devices have had retarded accpetance by the minerals exploration industry. Magnetic, electromagnetic, electrical, and gravity measurements took many years of development prior to their incorporation into geological exploration programs. All of the above techniques can be utilized for the direct identification of ores, but in practical usage, magnetics and gravity surveys outline the subsurface geology, while electromagnetic and electrical programs are aimed at the direct discovery of massive sulphides, graphite, and disseminated sulphides.

Remote sensing from space will not lead directly to the discovery of ores in North America, but will certainly become one of our greatest tools for outlining the geology which controls the location of ore deposits. In contrast, prior to 1960 Western Australia was a virgin prospecting area for iron ore and nickel sulphide deposits Remote sensing from space could have depicted the extent of the Hammersley iron formations and furthermore, could have outlined the gun barrel blue of the enriched ores now being mined as the largest open pit deposits in the world. The serpentine environment of the nickel sulphide deposit could have been outlined quickly by space imagery.

However, we explorationists must direct our mine finding effort to North America and, in particular, the United States. A new generation of

ore deposits remains to be discovered which are sub-alluvium. All of these ore deposits formed concomitantly with the enclosing rocks and therefore the mapping of controlling environment is the key to success.

Early utilization of remote sensing from space attempted to outline alteration areas and lineaments as direct guides to ore deposits. Lineaments are post ore. Alteration areas are readily identified by aerial photography. Therefore, early utilization has failed.

The future of remote sensing from space in the field of mineral exploration in North America will be its ability to assist the geologist in outlining preconceived ore controls. Rhyolite volcanic centers in eugeosynclinal systems and sub-volcanic magmatic patterns in porphyry systems are within the capability of the existing space sensing devices. The ability to view a large region, completely in scale, of a specific spectral band will be a major tool which will be of immense value to the explorationist. To accomplish this goal, satellites should be designed to measure the physical and/or chemical parameters of the ore controls for the various classes of ore deposits. The ore parameters can be defined up to certain limits. Therefore, meetings should commence between space technologists of image processing technology and earth bound explorationists.

ACTUAL USER EXPERIENCE AND NEEDS

Jon W. Davidson

The Superior Oil Company
Houston, Texas

ABSTRACT

Most of our present applications of Landsat digital data to geologic problems are limited by two parameters in the basic data:

1) Limited dynamic range of the present recording system.

2) Spectral position of the present bands.

Although the system records 64 gray levels, our experience indicates that a majority of the time 99% of the data are recorded in less than 20 levels. This limited dynamic range does not allow adequate delineation of the subtle tonal changes frequently associated with geologic features. The limited dynamic range is a particular problem in

ratioing, a technique we think is very useful in differentiating subtle geologic features.

The spectral position of the present LANDSAT bands were chosen to maximize discrimination of vegetation for the agricultural community. Unfortunately for the geologic community, the bands in Landsat-C and Landsat-D have been chosen with vegetation discrimination as their prime objective. In our experience, the most productive geologic work has been accomplished in nonvegetated areas and with the aid of ratios. The addition of a band in Landsat-D centered at 2.15 microns would be extremely useful to the geologic community.

THE CASE FOR STEREO

William V. Trollinger

Trollinger Geological Associates, Inc.
Denver, Colorado

ABSTRACT

The value of Landsat for petroleum and mineral exploration has exceeded the expectations of many users. However, Landsat has not yet had the striking impact needed in the fact of the growing worldwide energy crisis. Although Landsats C and D, as designed, have important improvements, they will still be deficient in certain aspects that limit their value for petroleum/mineral exploration.

One of the greatest needs is the capability for stereo viewing. The following comments are in defense of that statement.

Remote sensing for petroleum/mineral exploration is not new. The science of photogeology and its application to geologic mapping was made practical by research done on stereo viewing of aerial photography during World War II. Explorationists originally scoffed at the idea, but as the technique became more refined, they began to realize the value of the vertical, three-dimensional stereoscopic perspective for geologic observations. Photogeology is now an accepted and fundamental tool for petroleum and mineral exploration.

The greatest contribution photogeology has made for petroleum exploration is toward improved and efficient surface geologic mapping. And, the geologic map is the first and foremost tool of the oil explorationist!

The geologic map contains at least two important types of information; i.e., (1) stratigraphic (lithologic) and (2) structural. Differentiation of the various surface rock units and mapping of surface structural features is "state of the art" using stereo vision and quality imagery. However, these tasks are virtually impossible to accomplish without stereo, or in a monoscopic mode. This is because virtually all of the geological phenomena sought after by the photogeologist are three-dimensional features that are only partially displayed on the two-dimensional format of a single image. Their significance is clouded because they are perceived in only two dimensions, while in fact they represent profound three-dimensional elements to the geologic puzzle. Their correct interpretation as viewed at the surface holds the key for determining the character of the subsurface conditions and to unlocking the chronology of geologic events responsible for the emplacement of petroleum and mineral deposits in the subject area.

The space program offers the greatest potential for helping man rapidly locate and harvest the world's remaining hidden energy resources. To achieve this goal, however, the systems must be redesigned to better meet the needs of the exploration geologists. One of the first of these needs is stereo.

(Note: The paper of Keith Fullenwider, Sr., will be supplied at the Conference.)

IMAGE DATA AT THE NATIONAL SPACE SCIENCE DATA CENTER

James I. Vette

National Space Science Data Center, Code 601
Greenbelt, Maryland 20771

ABSTRACT

Many satellite instruments have produced image data in a variety of wavelengths ranging from microwaves to soft x-rays. These observations have been made mainly on the sun, Mercury, Venus, the earth, the moon, Mars, and Jupiter although other celestial objects and some satellites of the planets have been viewed. Many of these data, either in the form of photographic images or magnetic tape, have been deposited in the National Space Science Data Center (NSSDC) for distribution to interested users. The types of data and characteristics of the instruments which produced them are briefly described and additional documentation provided by NSSDC is cited.

I. INTRODUCTION

Much of the data obtained from satellites, particularly that from space science experiments, is available to users through the National Space Science Data Center (NSSDC) or its collocated international arm, World Data Center A for Rockets and Satellites (WDC-A-R&S). A significant fraction of this data is deposited in the form of images or digital magnetic tapes from imaging type instruments. In the past eight years, approximately sixty-two percent of the requests for data that NSSDC/WDC-A-R&S has satisfied have involved images on photographic film. And of this sixty-two percent, forty-five percent has been lunar photography, eleven percent has been planetary, and six percent has been of the earth.

In the framework of this symposium, the Data Center is a source of a diverse variety of image data; it is not involved in the development of image processing techniques nor is much analysis of this particular type of data carried out. However, the quality reproduction of the many types of photographic data and the cataloging of these images on microform represent an important effort of the center. In addition NSSDC is able to provide the digital data in a variety of formats, one of which is generally compatible with the computer facilities of the user.

This paper will describe the image data available through NSSDC/WDC-A-R&S and will also point out those areas where data are available only from other sources. Since requests for all types of data involving satellites come to the Data Center, it must serve as a switching center for such requests, routing them to the proper facility. It seems most useful to organize the discussion by the body of the solar system which is being observed. Remarks will be confined to the sun, Mercury, Venus, the earth, the moon, Mars, and Jupiter although observations of other celestial bodies and regions, as well as some satellites of the planets, appear in some of the data. The wavelengths range from microwave observations of the earth to x-ray images of the sun. Although the main focus of NSSDC is on extra-terrestrial observations, some earth observations are also handled.

II SOLAR OBSERVATIONS

The main observations of the sun by satellite imaging experiments have been, quite naturally, in wavelength regions which are inaccessible from the surface of the earth. Consequently, the solar imaging data available at NSSDC lies mainly in the ultraviolet and x-ray regions of the electromagnetic spectrum. There are some white light and Hα observations from certain Skylab experiments.

The earliest solar imaging data deposited was that from the OSO-4 Soft X-ray Telescope. This experiment obtained observations with a resolution of 1 arc-min in the four wavelength bands: 0.3-1.3, 0.3-2.1, 0.3-2.0, and 4.4-7.0 nm. A complete raster scan producing a 48- by 40- word array of x-ray intensities was obtained every 307s. This data is only available on digital magnetic tape since none of the images have been submitted for distribution.

The OSO-6 EUV Spectroheliometer obtained a large number of scans of the solar disk during its operational life from August 1969 to May 1970. The instrument covered the spectral range from 28-138 nm in 0.01 nm steps and had an angular field of view of 35 arc-s by 35 arc-s. Spectral scans were obtained by pointing the instrument to one of the positions in a 128- by 128-points reference grid centered on the solar disk; 16 minutes were required to obtain these scans. The full disk spectroheliograms were obtained at selected wavelengths with a raster containing 64 rows spaced at constant elevation increments of 42 arc-s with 96 points in each row. The scans required 8 minutes to complete. Sections of the solar disk 7 arc-min. in azimuth and 7.5 arc-min. in elevation were scanned in a small raster pattern consisting of 16 rows spaced 28 arc-min. apart with 24 points in each row. This small raster required 31 seconds to complete. This data is only available on digital magnetic tape.

Image data obtained on several Apollo Telescope Mount (ATM) experiments on Skylab are now available at NSSDC. The White Light Coronograph made observations in the 400-600 nm band with a net angular resolution of 8.2 arc-s of the solar corona extending from 1.5 to 6 solar radii. An example of the film data available from this experiment is shown in Fig. 1. The bright region at the center of the picture is from a calibration wedge illuminated by the sun which gives intensities from 10^{-8} to 10^{-10} that of the solar disk. The matrix of dots contains the information on the operating mode, filter wheel position, exposure, time of year, spacecraft roll, pointing error sensor, and internal alignment sensor. An extensive "Scientific Data Analysis Guide for the High Altitude Observatory White Light Coronagraph Experiment for the Apollo Telescope Mount" was prepared for NSSDC by the experiment team to aid users of the data.

The Skylab X-ray Spectrographic Telescope obtained full disk pictures in the 0.2-6 nm band
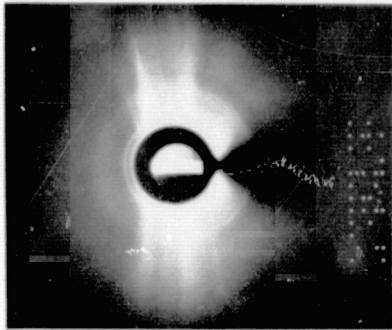
Fig. 1 Photograph of the solar corona taken by the Skylab White Light Coronagraph. This nine second exposure was taken at 1318 UT on August 22, 1973.

with a spatial resolution of 2 arc-s. Each x-ray picture is accompanied by a white light photo of the sun. A series of exposures were taken; a data set has been formed at NSSDC with the optimum exposure selected. An example of this data is shown in Fig. 2. The experiment team has produced a "Data Analysis Guide to the ATM S-054 X-ray Spectrographic Telescope" which is available through NSSDC.
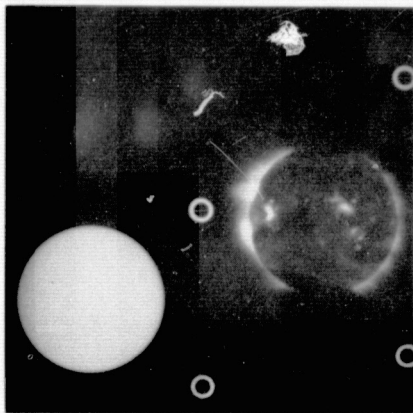

Fig. 2 X-ray and white light photograph of the sun taken by Skylab X-ray Spectrographic Telescope.

Data are also available from two Harvard College Observatory (HCO) experiments on Skylab. The experiments and their data are described by W. Harby and E. M. Reeves in a "Data Handbook for the Harvard EUV Spectroheliometer on ATM." The spectroheliometer operated in the spectral range 29.6- 135 nm with a 0.15 nm resolution. A raster scanning mode produced images of a 5.5 arc-min. by 5 arc-min. area of the solar disk in seven different wavelengths simultaneously. A mirror line scanning mode was used to acquire data in the seven spectral bands over an area of 5 arc-s by 5.5 arc-min. A spectral scan of a 5 arc-s by 5 arc-s area covering 5270 points in 3.8 min. was also obtained. The image data from this experiment are available on magnetic tape and as microfiche images of the raster scan and line scan using a 64 level gray scale. The raster picture consists of 60 lines with 120 elements per line while the mirror line scan contains 60 lines with 7200 elements. An auxiliary HCO experiment was two Hα telescopes used by the Skylab crew as pointing instruments. A film camera on one of these telescopes provided Hα images which are available on 35 mm film.

Data from other Skylab x-ray and EUV experiments are in the process of being deposited and will be available for distribution at a later date.

III MARINER 10 OBSERVATIONS OF MERCURY AND VENUS

Mariner 10, which had three encounters with Mercury and one with Venus, was equipped with two vidicon cameras capable of taking wide and narrow angle photographs of these two planets. Each camera had an eight position filter wheel and the optical systems to produce both an 11 x 14 degree (wide angle) and a 0.37 by 0.48 degree (narrow angle) field of view. There were 7,816 pictures of Venus and 5,514 pictures of Mercury obtained on this mission. The pictures contained 700 lines with about 900 elements per line.

These pictures were processed in several ways at the Jet Propulsion Laboratory (JPL) by the Mission Test Video System (MTVS) and by various digital processing techniques by the Image Processing Laboratory (IPL). Consequently for a given frame there are many different versions. NSSDC has received from JPL all three MTVS versions but only the "best" IPL version for each frame. Various IPL techniques are employed to achieve a desired result and it is only this selected image that is available from NSSDC.

The three MTVS versions are: (a) raw picture with contrast enhancement, (b) high-pass filtered picture, (c) vertical AGC. The raw pictures display the full range of brightness present in the original scene with the relative brightness between areas accurately preserved. The high-pass filtered version has had the variations in average brightness level reduced along each TV line by a high-pass filter. This allows for use of extreme contrast stretch to display maximum surface details without encountering a washout of high-and low-brightness areas. This version distorts the albedo. The vertical AGC version has been subjected to filtering along the vertical sets of picture elements. Each element is adjusted by a function of the average value of the picture elements in the column above the element being corrected. Examples of the three versions for a scene from Mercury are shown in Fig. 3. The data block to the right of each picture gives the supporting information about the original picture and the processing that was done.

Besides the MTVS and IPL pictures, indexes on 16 mm film, supplementary experiment data record on 16 mm film and picture catalogs on microfiche for each encounter are available. More details are available in "Status of Availability of Mariner 10 TV Picture Data", Data Announcement Bulletin (DAB) NSSDC 75-18, October 1975.

IV EARTH OBSERVATIONS

There have been a great many observations of the earth by satellite sensors, certainly more than any other planet with the moon being second. Because there are several different organizations that distribute the wide variety of earth observations to users, Table 1 has been included to clarify what each group handles. In keeping with the previous sections, only the holdings of NSSDC will be detailed here.

The Scanning Radiometer on Tiros 2 and 3 made measurements in the five spectral bands: 0.2-6.0 (reflected solar radiation), 0.55-0.75 (response of TV camera), 6.0-6.5 (water vapor absorption), 7.5-30 (terrestrial radiation) and 8.0-12.0 μm (atmospheric window). The terrestrial radiation channel

was eliminated on Tiros 4, while a 14.8-
15.5 μm channel (CO$_2$ absorption) was used
on Tiros 7 in place of the water vapor
absorption channel. These data are only
available on magnetic tape. Catalogs and
Users Guides for all four of these Tiros
satellites are available.

The Nimbus 1 and 2 satellite carried
a High Resolution Infrared Radiometer
(HRIR) which measured cloud cover and
terrain and cloud top temperatures in the
3.5-4.1 μm band. Besides the radiation
data on magnetic tapes there are 70 mm
film strips gridded with geographic co-
ordinates; one picture covers a swath
that runs nearby from pole to pole. In
addition, Nimbus 2 carried a Medium Re-
solution Infrared Radiometer (MRIR)
which covered the five bands 0.2-4.0,
5-30, 6.4-6.9, 10-11, and 14-16 μm.
There are 101 mm x 127 mm (4 in. x 5 in.)
photographic sheets available for each
band covering the complete daylight posi-
tion of the orbit. The photographs have
a latitude grid, time, and a gray scale
indicating temperatures. Nimbus 3 carried
an HRIR with two channels (0.7-1.3 and
3.4-4.2 μm) and an MRIR with five channels
(0.2-4.0, 6.5-7.0, 10-11, 14.5-15.5, and
20-23 μm). The format of the data is the
same as Nimbus 2. A new radiometer called
Temperature-Humidity Infrared Radiometer
(THIR) was flown on Nimbus 4 and 5 to
measure radiation in the 6.5-7.0 and 10.5-
12.5 μm bands with a resolution of 6.7 km.
A 70 mm filmstrip from the Nimbus 5 THIR
is shown in Fig. 4 for a nightime swath.
The grid point array makes lines for every
10° latitude and longitude, with points
spaced at 2° intervals along each 10°
line between 60° N and 60° S latitude.
Outside of these latitudes, the longitude
lines are every 20° with the points
spaced at 5° intervals while the lati-
tude markings do not change. A ten-step
calibration grey scale wedge provides the
approximate corresponding blackbody tem-
peratures on the film strips.

NSSDC also has data from the Nimbus
5 Electrically Scanning Microwave Radio-
meter (ESMR) which viewed the earth in
the 19.225-19.475 GH$_z$ band. The beam
width is 1.4° x 1.4° near the radii pro-
ducing 25 km resolution. An example of
the image format on 70 mm film is shown
in Fig. 5. The data orbit number iden-
tifies only the last data orbit; on most
frames two different orbits of data are
shown. At the left of the figure is an
11-step gray scale which serves to define
the ESMR brightness temperature on the
three swaths shown to the left of center
(nighttime) and the three swaths to the
right of center (daytime). The scale
changes for each numbered swath and is approximately
in the range 110-200, 180-262, 253-280 for swaths
1, 2, and 3, respectively. The time code in hours
and minutes UT is adjacent to the gray scale; the
top number corresponds to the left group of three
swaths and the timebars spaced at five minute inter-
vals. The geographic grids are similar to those in
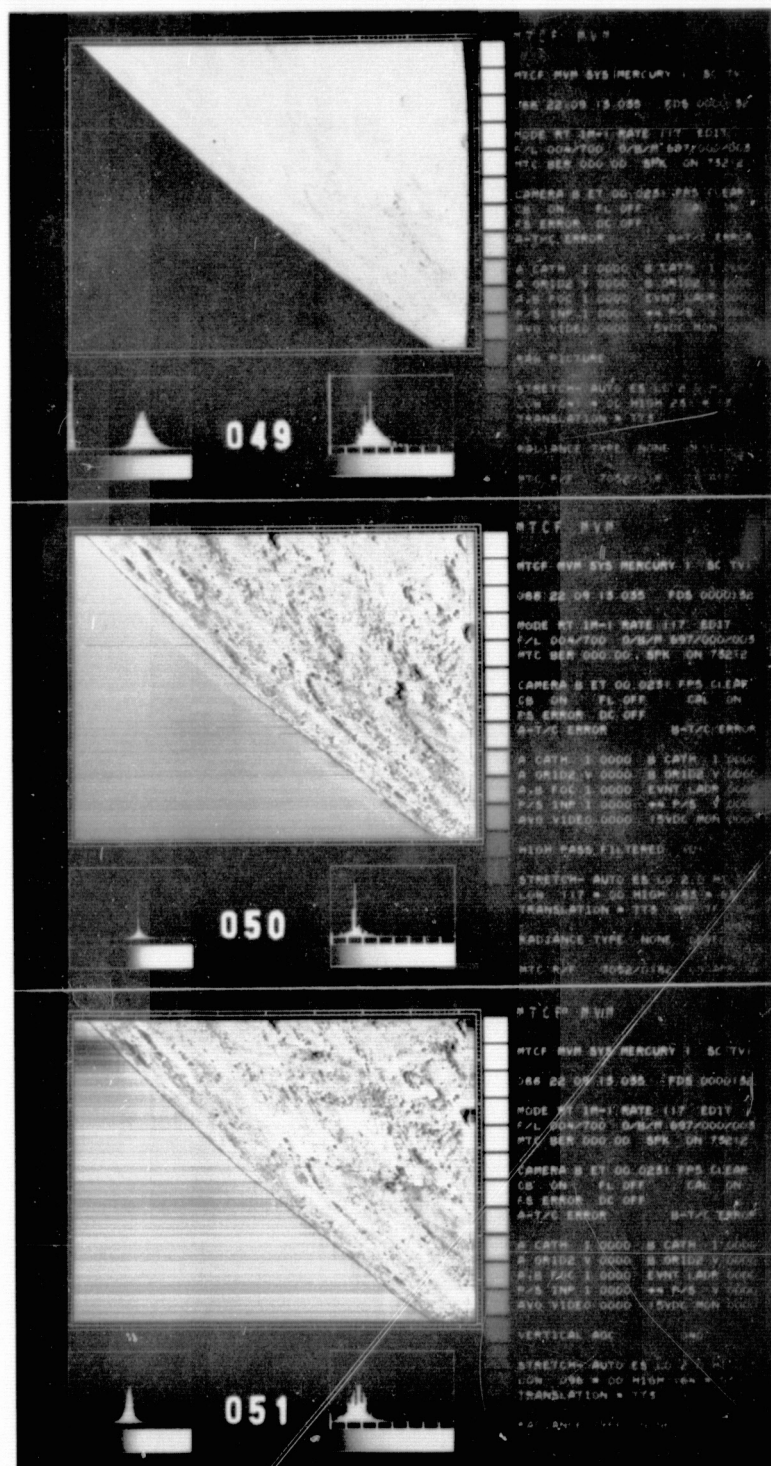
Fig. 3 The three differently processed MTVS versions of a
scene from Mercury taken by Mariner 10. Frame 049 is the
raw picture, frame 050 is the high-pass filtered, and frame
051 is the vertical AGC.

Fig. 4 but the equator, North Pole and South Pole
are indicated with appropriate abbreviations.

There are detailed Nimbus 1-6 User's Guides
which describe all the experiments, discuss data
processing, and indicate where data will be archiv-
ed. In addition, there are numerous catalogs pub-
lished for each satellite which indicate just what

Distributors of Satellite Earth Imaging Data

Satellite	Experiments	Distributor
Apollo 9	Multispectral Terrain Photography	TAC*
Apollo Series	Visible & IR Pictures of Earth	TAC
Apollo 16	Far Ultraviolet Camera	NSSDC
ATS 1 & 3	Spin Scan Cloud Camera	SDSB*
DMSP Series	Auroral Imagery from Scanning Radiometer	NGSDC*, NSSDC
ESSA 1-9	TV Camera	SDSB
Gemini 3-12	Synoptic Terrain Photography	EROS*
ISIS 2	Auroral Photometers	NSSDC
LANDSAT 1 & 2	Multispectral Scanner, Vidicon Camera	EROS, SDSB
Nimbus 1 & 2	Vidicon Camera	SDSB
Nimbus 1-5	Scanning Radiometers	NSSDC
NOAA 1-5	Visible & IR Scanning Radiometers	SDSB
SMS 1 & 2, GOES 1	Visible Spin Scan Radiometer	SDSB
Tiros 1-10	TV Camera	SDSB
Tiros 2, 3, 4, 7	IR Radiometers	NSSDC

*EROS - EROS Data Center, Sioux Falls, SD
 NGSDC - National Geophysics & Space Data Center, Environmental Data Services (EDS),
         NOAA, Boulder, CO
 SDSB - Satellite Data Services Branch, National Climatic Center, EDS, NOAA, Washington, D.C.
 TAC   - Technical Applications Center, U. of New Mexico, Albuquerque, NM

Fig. 4 Nimbus 5 THIR nighttime swath from the 11.5 μm channel. The data began at 1615 UT on March 12, 1975, and the longitude values of 93, 107, 115, 123, and 127E are marked at the right hand side of the strip below the 60S, 30S, 0, 30N, and 60N labels, respectively.

Fig. 5 Nimbus 5 ESMR swaths at 0025 UT on July 31, 1974.

Fig. 6. The geographic positions can be determined from ephemeris information and coordinate grids which are available; automated gridding such as displayed for Nimbus data is not done for this data. The auroral films are made by NGSDC (see Table 1) from the DMSP originals and copies are made available to NSSDC. A DMSP User's Guide has been published by the Air Weather Service United States Air Force in December 1974 (AWS-TR-74-250).

data is available and shows many of the film images. These documents are all available from NSSDC. All of Nimbus data discussed earlier in this section are available on digital magnetic tape.

The Defense Meteorological Satellite Program has had several satellites, DMSP, which carry two scanning radiometers each sensitive in the 0.4-1.1 and 8-13 μm range. The two differ in resolution by a factor of six, with the highest resolution being 0.6 km at the radii. Polar night passes containing auroral imagery have been released on 35 mm film. An example of such imagery is shown in

There were two auroral photometers flown on ISIS 2. One was a scanning photometer sensitive at 558.1 ± 0.9 and 391.5 ± 1.3 nm using interference filters. The 5° wide earth scan produced by the satellite spin was divided by an image dissator into 0.4° by 0.4° regions. The other photometer was sensitive at 630.0 ± 1.2 nm and had a white light

channel. Both sensors had a 2.5° circular field of view. The intensity maps of these two photometers are available on a single magnetic tape.

The final imagery taken of the earth that we discuss here is that from the Far UV Camera/Spectrograph flown on Apollo 16 and operated on the lunar surface during the astronauts' visit. Besides photographing certain nebula and other celestial objects, the earth's atmosphere and the geocorona were viewed. The experiment used a 75 mm electronographic Schmidt camera with a potassium bromide cathode and a 35 mm film magazine and transport. Photographic images were obtained in the 105-155 and 125-155 nm bands. Intensity contour maps on 16 mm film are available.
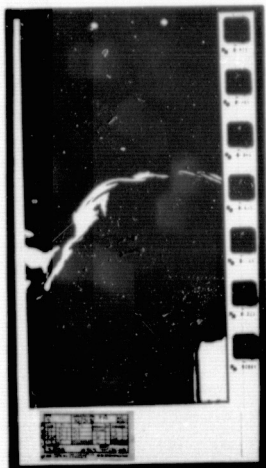
Fig. 6 DMSP auroral photography taken at 1046 UT on Jan. 5, 1976. The data block at the bottom allows ephemeris data and geographic grid to be selected.

V LUNAR IMAGES-RANGER THROUGH APOLLO

There are a number of different satellite series which have photographed the moon in a variety of ways. The first close-up pictures were obtained by Ranger 7, 8, and 9 prior to their impact on the lunar surface. This was followed by the Surveyor series, the Lunar Orbiter series, and finally the Apollo flybys and landings. In addition, there are some photographs available from the Soviet satellites Luna 3, 9, and 13 as well as Zond 3 and 8.

The TV camera systems on the final three Rangers were identical. There were two wide angle and four narrow angle cameras and two sets of channels, a partial and a full. The full channel contained one wide and one narrow angle camera. These full channel picutres contain 1150 lines while the partial channel ones contain 300 lines. A more detailed description is given in "Ranger 7, 8, 9 TV Cameras" Data Users Note (DUN) NSSDC 68-06 which contains ordering information.

Surveyors 1, 3, 5, 6, and 7 were successful in landing and operating on the lunar surface. A TV camera with resolution down to one mm was on each spacecraft. All except Surveyor 7 provided data from selected mare sites for the Apollo Program; Surveyor 7 obtained data in the lunar highlands. The slow scan vidicon camera had both a 200 line and a 600 line scan. NSSDC has 70 mm film generated from the original video data, mosiacs, and enhanced photographs. The following documents are pertinent for ordering Surveyor pictures: (a) "Surveyor 1 Lunar TV," DUN NSSDC 67-30, (b) "Catalog of Surveyor 1 Television Pictures," NSSDC 68-10, and (c) "Surveyor 5 Catalog" NASA SP-341.

The Lunar Orbiter Series had five spacecraft to accomplish the task of photographing the moon within 5° of the lunar equator down to a one meter resolution. Since this was accomplished in the first three missions, Lunar Orbiter 4 and 5 were

placed in polar orbits with the latter giving full coverage of the farside of the moon. Some 99% of the moon was photographed with a resolution of 60m. The Lunar Orbiter camera was a dual lens one in which the film was developed and then scanned to transmit data back to ground stations where the image was reconstructed. Both high and medium resolution frames were taken and image motion compensation was employed. The reconstructed pictures consisted of 35 mm framelets; a medium resolution frame consisted of 26 framelets while the complete high resolution picture consisted of 86 framelets - three frames with some overlap. Although there have been enhancements of various pictures by certain groups, all of these have been superceded by the Langley Research Center enhancements. These are in the form of 20" x 24" negatives. There is a microfiche catalog made by Cal Tech which contains all the enhanced photographs. More details and ordering information are available in the documents: (a) "Lunar Orbiter Photographic Data," DUN NSSDC 69-05 and "Status of Availability of Lunar Orbiter TV Picture Data," DAB NSSDC/WDC-A-R&S 76-02.

There are a limited number of pictures available from the Luna and Zond series. Luna 3 and Zond 3 were lunar flybys. Zond 8 flew around the moon and returned to earth while Luna 9 and 13 were landers. The photos that NSSDC has consist of: (a) thirty Luna 3 pictures from "Atlas of the Farside of the Moon," (b) three Luna 3 pictures from "First Photographs of the Farside of the Moon" (c) twenty-three Zond 3 pictures from "Atlas of the Farside of the Moon Part 2", (d) nine Zond 8 photos provided as 4" x 5" prints, (e) twelve unidentified photos from Zond 6, 7, & 8, (f) Luna 9 & 13 photos in "First Panoramas of the Lunar Surface."

More details on Zond 8 are available in "Status of Availability of Zond 8 Lunar Photography" DAB NSSDC/WDC-A-R&S 76-01. A Zond 8 photograph is shown in Fig. 7.

Finally, the most extensive set of pictures of the moon come from the Apollo series of spacecraft. Apollos 8, 10, 11, 12, 13, and 14 obtained photos with Hasselblad 70 mm cameras and 16 mm motion pictures with Mauer cameras. In addition, Apollo 12 had a 35 mm close-up stereoscopic camera and Apollo 14 carried a Hycon Topographic camera. Apollo 15, 16, and 17 produced the largest number and variety of photographs. To indicate this variety we list in Table 2 the photographic equipment used on Apollo 17 with some of the characteristics of each camera. The details of all the Apollo photography can be obtained from NSSDC DUN's and/or catalogs for each Apollo Mission. Two samples of Apollo 17 photography, the metric camera photograph in Fig. 8 and the panoramic camera reduced photo in Fig. 9, are shown.

VI MARS

The first pictures of Mars were obtained from Mariner 4 with a TV camera that had a 1.05°-by 1.05°- field of view. The vidicon output underwent analog to digital conversion with each picture containing 240,000 bits; 200 lines with 200 elements per line and 6 bit intensity level. Twenty-one complete and one partial pictures were returned from this first photographic mission. The Data Center has several different versions of these pictures available: (a) air brush renditions-interpretation of what the surface of Mars may look like, (b) calibrated and geometrically corrected version
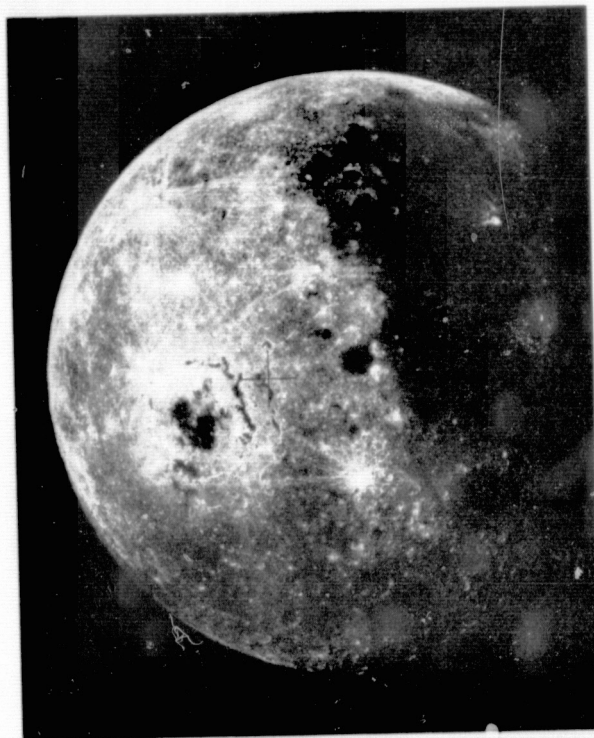
Fig. 7 Zond 8 photograph taken on October 24, 1970. The sunset terminator is at 20°W long. Oceanus Procellarum is dark area in upper right with the crater Aristarchus standing out brightly. Dark crater at center is Grimaldi. Mare Humorum is down right of center and Mare Orientale is left of center. The main part of the lunar backside occupies the left section of the photograph.

Fig. 8 Apollo 17 mapping (metric) camera photograph showing the Apollo 17 land site. Site is marked by arrow. Frame AS17-2309.

Table 2. Summary of Apollo 17 Primary Photographic Equipment

| Camera	Focal Length(mm)	f/no.	Image Size (mm)	Field of View (deg)
		Service Module		
Mapping	76	4.5	115 x 115	74 x 74
Stellar	76	2.8	24 x 32	18(ver x 24(hor)
Panoramic	610	3.5	115 x 1150	11(fore, aft) x 108(across)
		Command Module		
Hasselblad	80	2.8	52 x 52	37.9(side), 51.8(diag.)
	250	5.6	52 x 52	12.5(side), 17.6(diag.)
Maurer, Motion	10	1-22	8 x 10	54.9(hor) x 41.1 (ver)
	18	1-22	8 x 10	32.6(hor) x 23.4 (ver)
	75	2.4-22	8 x 10	7.9 (hor) x 5.7 (ver)
	229	8	8 x 10	2.1 circ.
Nikon	55	1.2	24 x 36	36.0(hor) x 24.0(ver)
Westinghouse TV	–	4-44	8 x 10	–
		Lunar Surface and Lunar Module		
Hasselblad	60	11	52 x 52	46.9(side), 63.4(diag)
	500	11	52 x 52	6-2(side), 8.8(diag)
Maurer, Motion	10	1-22	8 x 10	54.9(hor) x 41.1(ver)
RCA TV	12.5-75	2.2-22	8 x 10	–
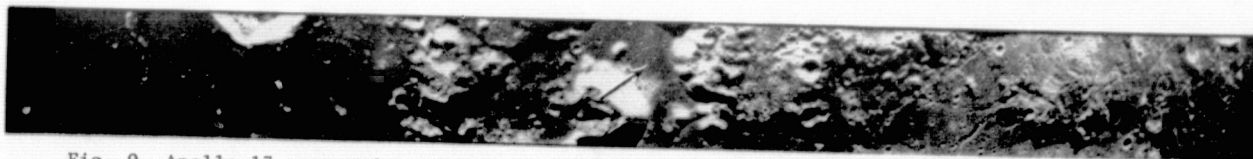
Fig. 9 Apollo 17 panoramic camera reduced photograph showing the Apollo 17 landing site.

enhanced in contrast, (c) a version similar to (b) but also sharpened, (d) a negative version of the original picture format, (e) a calibrated contrast-enhanced version in the original picture format, (f) a version similar to (e) but also sharpened, (g) a fluctuation plot in which smooth areas are rendered as dark and locally rough areas are rendered as light.

Mariners 6 and 7 each had two TV cameras, one wide angle (11° x 14°) and one narrow angle (1.1° x 1.4°). The wide angle camera was equipped with red, green, and blue filters which rotated in sequence. The narrow angle camera had a minus blue haze filter. In the near encounter mode, the camera shutters operated alternately and were timed to provide overlapping of the wide angle pictures, with the narrow angle picture falling inside the

spacecraft. Resolution of 500 meters and 50 meters were obtained by the low and high resolution cameras, respectively, at the periapsis altitude of 2000 km. More than 7300 pictures were obtained of the Martian surface, the Martian satellites, and celestial objects. A variety of picture enhancement techniques have been applied to the original data resulting in more than 30,000 photographs being available through NSSDC. Besides the three MTVS versions mentioned for Mariner 10, there are a wider variety of IPL versions available. The two main versions are albedo and maximum descrimination. In addition to these the Data Center has microfiche catalogs of the MTVS and selected IPL versions, the JPL mosaics, as well as supporting data and indexes on microfilm or microfiche. More details are available in "Status of Availability of Mariner 9 TV
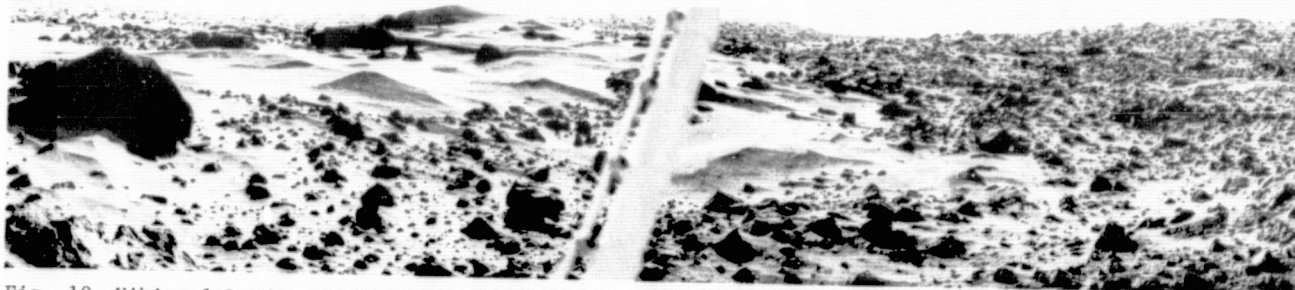

Fig. 10 Viking 1 Lander photograph showing a dune field similar to many seen in the deserts of the earth. This picture was taken on August 3, 1976, at 0730 Mars local time. The Viking meteorology boom cuts through the center of the picture.

overlapped portion. There were 25 near encounter and 50 far encounter pictures from Mariner 6 and 33 near encounter and 93 far encounter pictures from Mariner 7. The pictures contained 704 lines with 945 elements per line.

The image data have been processed in several ways by JPL to bring out the details that appear subtle in the raw-analog version. The following versions are available on 70 mm black and white film: (a) raw analog, (b) maximum discrimmability (optimal presentation)-High pass filtering, small scale enhancement, system noise removal and geometric corrections, (c) photometrically decalibrated output signal corrected digitally to actual scene luminance, (d) maximum discriminability, alternate contrast enhancements-similar to (b) with some differences.

In addition to these products, photo mosaics of near encounter photographs have been made. A more detailed description of the available products can be found in "Mariner 6 and 7 Photographic Data", DUN NSSDC 71-09.

Mariner 9, which was operational in orbit around Mars from November 14, 1971, to October 27, 1972, obtained the largest number of pictures of the planet taken to date by a single satellite. The TV photography experiment was similar to that flown on Mariner 6 and 7 and was operated in the near encounter mode described for these earlier
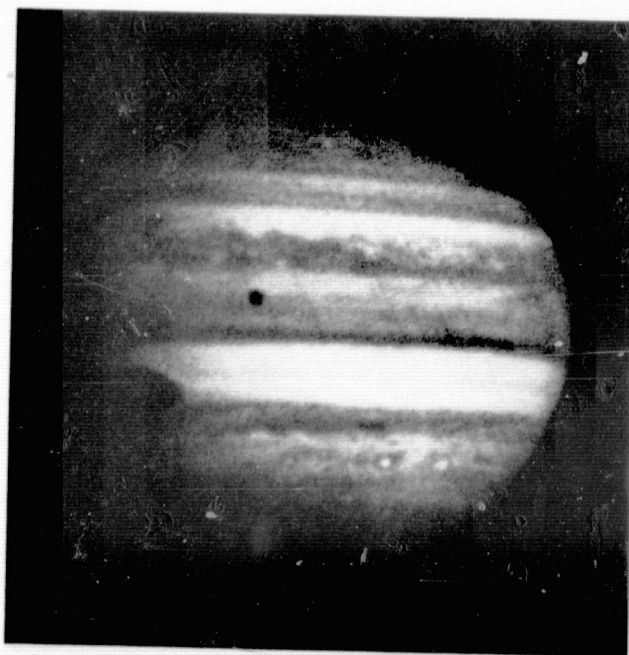

Fig. 11 Pioneer 10 photograph of Jupiter showing the bands, the Great Red Spot and the moon, Io, crossing the Jovian disk. This picture was taken on December 2, 1973, at 0656 UT when the satellite was 2,6000,000 km from the planet.

Picture Data," DAB NSSDC 74-05 and in "Catalog of
Ionospheric Data," NSSDC 75-07.

At the time of this writing the Viking 1 and 2
orbiter pictures are just starting to arrive at the
Data Center. It will be several months before in-
dexes and supporting information are available and
an appropriate catalog or DAB can be made. A very
recent picture taken by the Viking 1 lander showing
the sand dunes is shown in Fig. 10.

VII JUPITER

The Imaging Photopolarimeter Experiment flown
on Pioneer 10 and 11 has provided the only satel-
lite photographs of Jupiter. Some of these frames
include the Jovian moons, Io, Europa, and Ganymede.
The instrument produced spin scan images in two
wavelength bands, blue (390-490 nm) and red (580-
700 nm), with maximum resolution of 200 km. Black
and white negatives either 127 mm x 178 mm (5 in x
7 in) or 203 mm x 254 mm (8 in x 10 in), depending
on the frame number, for both the blue and the red
images have been deposited for Pioneer 10. An ex-
ample of a blue image showing the whole disk of the
planet with Io visible is shown in Fig. 11. An in-
dex for these frames is available but no formal
documentation has been prepared yet.

Color photographs using a false green have
been produced for some frames of both Pioneer 10
and 11 and are currently on file. These photo-
graphs have been released to the press by the Pio-
neer Project Office as they have become available.

VIII CONCLUDING REMARKS

A great wealth of satellite image data in many
different formats and processed in a variety of
ways is available from NSSDC. A continuing inter-
est in meteorological observations and lunar sci-
ence coupled with a growing interest in comparative
planetology and in earth resources has made such
data be in high demand for the past eight years.
It appears that the useful lifetime of such data
will span at least two decades and the need for
active repositories, as an important link in opti-
mizing our scientific understanding of such data,
continues.

The image data discussed in this paper as well
as data from other types of experiments flown on
scientific satellites can be obtained by writing to:

    National Space Science Data Center
    Code 601
    NASA/Goddard Space Flight Center
    Greenbelt, MD   20771

    Telephone: (301) 982-6695

PANEL ABSTRACTS

"TODAY'S USER: HOW DO WE GIVE THEM WHAT THEY NEED?"

William B. Green, Chairman

Participants:

O. Glenn Smith

Stanley C. Freden

Darden W. Mooneyhan

James A. Dunne

Robert R. Jayroe, Jr.

Donald T. Lauer

Frederic C. Billingsley

TODAY'S USER: HOW DO WE GIVE THEM WHAT THEY NEED?

William B. Green

Manager, Science Data Analysis Section
Jet Propulsion Laboratory
Pasadena, California

About seven years ago when I arrived at JPL's Image Processing Laboratory, the world of computer image processing was quite different than it is today. The art of recording digital imagery, processing it with a computer, and recording a processed image on film was highly specialized. There were not a large number of systems producing digital imagery, and the number of film recording devices and display systems commercially available was rather limited. At that time, the display systems and film recorders we used were designed and often built by JPL. The Mariner 6 and 7 Mars missions had been completed, and the two spacecraft had returned the magnificent total of approximately two hundred digital images to earth. Our main customer was the television experiment team, a relatively small number of scientists who, I think, were quite frustrated with the amount of time it took to process the imagery and provide them with a usable product. Eventually, the experimenters got their complete sets of cleaned up imagery, the National Space Science Data Center got their photographic products, and the record of the processing required to produce the final products was retained at JPL in a single large notebook.

The world has changed substantially in the last few years. In the planetary program, we have made the transition from missions returning a couple of hundred images to the Viking mission returning imaging data from four spacecraft operating simultaneously and returning as much data in one day as the Mariner 6 spacecraft returned during its active lifetime. The Landsat satellites have produced thousands of multispectral images and have imaged the entire earth many times over. Radar imaging satellites and future Landsat satellites will produce data volumes that may dwarf the current Landsat return.

The large volume of data returned by NASA satellites during the past few years has had several effects. The most obvious is the development of image processing capabilities at many organizations. In addition, manufacturers have quickly identified a rapidly growing market, and a large variety of display and film recording hardware is now commercially available at reasonable cost. The growth in minicomputer capabilities and the large number of commercially available image display devices has significantly lowered the cost of acquiring an image processing capability.

Finally, Landsat has provided frequent, periodic global multispectral coverage at a time when man is attempting to assess the impact of his existence on his environment and also attempting to evaluate exactly how much of his expendable resources remain available to him for his own use and for future generations. The historically accidental merging of a technological capability with the new requirement for evaluation of complex factors on a global scale has increased dramatically the number of users of digital imagery. The planetary explorers have been joined by a growing number of diverse and demanding users seeking to deal with the fundamental issues of man's future on earth.

Despite the growing number of data users, the broad scope of the applications of digital image processing and the increase in data volume that must be accommodated, it seems to me that the techniques which must be used to "give the user what he needs" have not really changed that much since the days when a single notebook could hold all the processing records for an entire mission and a second notebook could hold all the image products produced on that mission. At JPL's Image Processing Laboratory, we have been the most successful when we have followed these guidelines:

1. Develop a trained staff with diverse backgrounds. One of our main strengths at the Image Processing Laboratory is the diversity of our staff. A facility devoted to computer processing of large volumes of scientific data obviously requires individuals with good computer science and programming background and a few individuals who can perform computer systems design and system analysis. In addition, we have found it essential to have individuals on our staff who are primarily trained in the disciplines represented by our users. My Section includes individuals with degrees in engineering, the physical sciences, geology, geography, astronomy, geophysics, chemistry, and biology. In order to interpret user requirements correctly, it is essential to have key individuals on your staff who can talk the user's language and at the same time understand the capabilities of your own facility. Outstanding programmers are required to insure successful and efficient implementation of user requirements; however, you may produce an elegant system that does not meet the user's requirements if your resident staff does not include individuals who can talk the user's language.

2. The users should be on-site at the implementing organization as much as possible. The best of all possible worlds as far as we are concerned is a world in which the ultimate user of our products has at least one staff member permanently resident at our facility. This assures maximum information exchange between implementer and users, and also results in a highly trained and sophisticated member of the user's staff. The resident user staff member can become the training and information focal point within his own organization. If the user agency cannot provide a full-time individual, the next best situation is to arrange for long term periodic visits by members of the user's staff. We have found that a single review and status meeting at, say, monthly intervals that lasts for half a day or even a couple of days is absolutely no substitute for the situation in which working members of the user's staff are actually resident within the implementing organization during requirements definition, design, and system development phases of a given project. And, of course, the extensive user involvement should begin as early as possible.

3. Appoint a member of your staff as a "user ombudsman." At an early date on each project designed to meet the needs of an outside user, you should name a single individual as a user representative, or user ombudsman. This individual should be selected from your staff and should be one conversant with both the user's technical areas of interest and the capabilities of your own facility and staff. As your employee, you should inform him that you will evaluate his performance based on the degree to which he successfully represents the user's needs and interests within your own facility. This keyindividual should become familiar with the user's requirements and objectives and should serve as the single interface point with the user for coordinating desired changes and modifications to the system design during implementation. In addition, for your own staff he can serve effectively in clarifying user requirements, resolving conflicts between requirements and capabilities, and performing systems tradeoffs as they are required. He should serve as a single source to which your staff refers all inquiries relating to user requirements and needs. In fact, if does a good job, your staff may ultimately forget that he is not actually employed by the user.

I feel that the best job we have done at the Image Processing Laboratory in following these guidelines was our experience with the Viking Lander project. The Lander Imaging Team had several full-time members resident at JPL several months prior to Mars landing who were housed within my Section, appeared on my Section telephone list, and were even active within the computer facility user group. The ultimate users of our imaging system on the Lander Imaging Team became so integrated into our facility operations during the mission that several members utilized our timesharing system to create and maintain their own private mission data files. At least two members of the Imaging Team actually became proficient at interactive image processing using terminals and displays in our computer user

areas alongside our own staff members. Early involvement by the users and the ability to have full-time resident members housed within our facility for long periods of time created a small nucleus of highly trained Imaging Team members who eventually assisted us in educating the remainder of the Imaging Team. In addition to the on-site Team members, two additional Team members assumed the responsibility for frequent and detailed communications regarding program and system development for the Viking Lander Imaging Team. The end result of our commitment to encourage frequent and in-depth contact with our ultimate users led to a successful implementation of the Viking Lander second order image processing system for flight operations. That implementation included replacement of our existing CPU; acquisition and integration of several new image display systems; design and development of a minicomputer based system for interfacing new dial-up terminals hard-lined graphics terminals, and image displays to the central CPU; modification of the complete IPL program library to enable successful operations under a completely new operating system and computer environment; development of a new executive language for interactive image processing and display; development and documentation of approximately thirty new programs for Viking applications; and implementation of an automated data base management system for cataloging and management of a large number of Viking Lander and Orbiter images on disk and tape. We were able to complete all of this activity on schedule and within budget. Even more significant to me, however, is the fact that the system satisfied the user requirements in actual flight operations. I feel that this is because we operated by the three guidelines listed above. Without a trained staff, early and in-depth involvement of the users within our facility during development and implementation, and a user ombudsman on our staff, I doubt if our completed system would have met the actual needs of our users during flight operations.

TODAY'S USER: HOW DO WE GIVE THEM WHAT THEY NEED?

O. Glenn Smith

Lyndon B. Johnson Space Center
Houston, Texas

ABSTRACT

I would like to discuss the differences between "what they want" and "what they need," and point out some pitfalls in selecting areas for NASA's work in developing applications.

We have both "problems looking for solutions" and "solutions looking for problems" in today's Earth Resources Program. For example, we have work in areas of future applications, such as microwave sensing of the earth's surface, which is not directed toward any specific problem, but is expected to contribute to parts of many problems in the future. On the other hand, we have projects aimed at some specific user's problem. These are always complicated, however, because the operational use of remotely sensed information is only a part of the problem facing the user. NASA cannot and should not attempt to solve the total problem of the user.

In selecting application areas for NASA work, we attempt to use such criteria as potential benefit, cost of system implementation, and chance for success. We also ask users what they need. I want to point out some pitfalls in the process.

If we ask operations people in the user community, we are likely to get requests for short term help, such as the use of our research aircraft to obtain photo coverage of a state or other area for real operational use. The problem here is that NASA's job is to develop future applications, not serve as a service organization to other agencies, or to compete with private industry.

When we ask individual researchers what they need, we sometimes can end up with a project that appears to us to have great promise, but when the research phase is complete and it is time to put the system into daily use, the operational people discover some practical problems that completely alter the expected result.

Of course the systems configuration, both satellite and ground, must support a variety of potential users and must represent a compromise.

Our conclusion is that we must recognize that the expected operational user must cope with integrating a remote sensing solution into the rest of his operating system. His advice to NASA is required, since NASA does not understand the non-remote sensing part of the problem. In addition, we must keep in mind that NASA will not be successful in any of these ventures until the users are successful in the operational application of Earth Resources data. As for the potential pitfalls, we believe that users must be directly involved in our applications projects — not just any users, but appropriate personnel in the user's organization. In addition to researchers, we prefer high-level management attention and the participation of the operational segment of the user's organization.

TODAY'S USER: HOW DO WE GIVE THEM WHAT THEY NEED?

Stanley C. Freden

Goddard Space Flight Center
Greenbelt, Maryland

ABSTRACT

Data requirements for the user community vary widely in terms of product types and timeliness. For example, near shore sea ice mapping can be accomplished with relatively low resolution black and white photographic products but the information is required within hours or days for routing ships. On the other extreme, high quality digital data are required for some geologic purposes but timeliness is not critical — weeks or months will often suffice.

To accommodate these data requirements, the Goddard Space Flight Center (GSFC) is implementing an entirely new processing facility for Landsat and other satellite imagery. Called the Image Processing Facility (IPF), it will preprocess all Landsat imagery in the digital domain; radiometric and geometric corrections will be applied to each scene. Up to 50 scenes will be stacked on each High Density Tape (HDT). The HDT's will become the archival product for further data production and distribution. Computer Compatible Tapes (CCT's) and film products will be subsequently productd from the HDT's. User products will consist of copies of the HDT's, CCT's, and film products at various scales from $1:10^6$ and larger.

Both the MSS and RBV data will be geometrically corrected digitally using ground control points. This will allow for accurate temporal registration of the data with no degradation of radiometric quality.

Turnaround time from acquisition of data at GSFC to distribution to the federal data centers will be reduced to 1-2 days compared to the present 2-4 weeks.

Further discussion and elaboration will be given of the data products and data product requirements as we presently see them for the Landsat-C time frame.

TODAY'S USER: HOW DO WE GIVE THEM WHAT THEY NEED?

Darden W. Mooneyhan

Lyndon B. Johnson Space Center
Earth Resources Laboratory
Slidell, Louisiana

ABSTRACT

Today's user of remote sensing technology is faced with a high rate of development that is increasing the lag between the developer and the user. Additionally, increased sophistication often requires expensive institutional changes and long term personnel training. The user, in order to justify the changes required to adapt to the new remote sensing technology needs:

(1) To be convinced that his specific problems can be solved

(2) To be assured that there is a reliable operational data source

(3) Training in the data handling and applications techniques to the extent that he is comfortable in the role of selecting hardware/software systems and/or services to solve his problems in an operational mode, and

(4) Cost effective data handling systems or data handling services to meet specific application needs.

These needs are being worked on in varying degrees with more emphasis on (1) in the past and with some increase in emphases on (2), (3), and (4) presently. It should be remembered that the easy part of technology transfer is for one technician to convince another that a new technology has application, the more difficult and time consuming part is convincing management, committees and legislatures to adopt a new technology into their institutions.

SEASAT-A DATA ACQUISITION AND DISTRIBUTION

James A. Dunne

Jet Propulsion Laboratory
Pasadena, California

ABSTRACT

Scheduled for launch in the second quarter of calendar year 1978, the Seasat-A satellite is designed to provide all-weather global monitoring of sea surface temperature, significant wave height, surface wind speec and direction and departures from the marine geoid corresponding to ocean dynamic processes. These data will be obtained from an array of microwave instruments, two active ones (short pulse radar altimeter and wind field scatterometer) and one passive (scanning multi-channel microwave radiometer). An experimental L-band synthetic aperture radar, operated on a selected basis for approximately four percent of the time, will provide land and ocean images with a resolution of 25 meters and a swath width close to 100 km for the study of coastal processes, sea ice and ocean wave characteristics. The mission objectives focus on an evaluation of the performance of the instruments in terms of their capabilities to characterize the desired geophysical quantities and the utility of such measurements to the study and exploitation of the world's oceans.

Data will be returned from the satellite in three separate streams: the 25 kbps real-time stream containing all non-SAR data, a playback stream (800 kbps) containing 25 kbps data which had been stored for later retransmission, and the 20 MHz analog SAR data stream, obtainable only in real time by a suitably equipped tracking station. Following receipt and recording at the tracking station, Seasat data is transmitted to the Goddard Space Flight Center, Greenbelt, MD, which in turn delivers to the Seasat Project at JPL sorted, merged and time-tagged telemetry data from the 25 and 800 kbps streams, unprocessed SAR raw data tapes, preprocessed tracking data, spacecraft ephemeredes and attitude solutions. Following the receipt of these data, the Pojrect data processing subsystem generates processed instrument, orbit and ancillary data for the experiment teams and other users. The SAR data is handled separately, being delivered in raw form to the SAR Data Processing facility at JPL, where the dispersed doppler and phase history data is decompressed and correlated

to form usable radar images. Further processing of these data is conducted by standard image processing techniques, both digital and optical. Outputs of the data system network will be provided to the user community through a number of channels, including normal user agency data archiving and distribution facilities, such as NOAA's Environmental Data Service, and the Geological Survey's EROS Data Center. Agreements for the conduct of these archiving and distribution functions are now being negotiated with the relevant user agencies.

Seasat-A was conceived as a precursor to a possible ocean monitoring satellite system by an advisory group consisting of NASA, academic, industrial and user agency scientists chaired by John R. Apel of NOAA. Economic benefit studies have indicated that an operational satellite system based upon the Seasat model could provide substantial savings in the conduct of a number of commercial and government-supported activities on the world's oceans and in coastal provinces. Evaluation of the performance of the Seasat-A instrument payload and the utility of the data they provide against the assumptions used as a basis for these studies will provide a firm foundation for future Seasat program plans.

Corresponding to each of the instruments, a team of NASA, academic, and user agency scientists has been formed for the purpose of evaluating their geophysical performances and end-system utility. The role of NASA personnel on these teams is generally a support one, with team direction and bulk membership being provided by the academic and user community. The teams place requirements on the satellite, data and mission operations systems in accordance with the scientific and applications experiments deemed necessary to provide the data upon which a thorough, objective evaluation can be made. They also stipulate the surface truth activities required, both pre- and postlaunch, to support algorithm development and the conduct of in-flight evaluation experiments and case studies.

MSFC/TENNESSEE STATE PLANNING OFFICE CLASSIFICATION

TECHNIQUE ASSESSMENT

Robert R. Jayroe, Jr.

George C. Marshall Space Flight Center
Marshall Space Flight Center, Alabama

ABSTRACT

The Tennessee State Planning Office (TSPO) of Nashville, Tennessee, and MSFC cooperated in an effort comparing classification results derived from Landsat imagery using several software programs with a land cover map obtained by interpreting high altitude aerial photography over the Bald Knob, Tennessee Quadrangle. The objectives of the effort were to acquire information on the feasibility of using Landsat imagery and classification software to obtain equivalent land cover maps and to acquaint TSPO with various classification software and their relative performances. The main results of the effort indicated that:

1. The feasibility was conditional.

2. Approximately 30% of the pixels occurred at a transition from one feature to another and accounted for nearly all of the misclassification.

3. If transitional pixel misclassification was ignored, there was essentially no difference between the various classification techniques.

4. The use of multi-temporal information increased the classification accuracy approximately 3%.

TODAY'S USERS: HOW DO WE GIVE THEM WHAT THEY NEED?

Donald T. Lauer

U.S. Geological Survey
EROS Data Center
Sioux Falls, South Dakota

ABSTRACT

The rate of development of remote sensing technology has increased much faster than the rate at which this same technology has been accepted and used in an operational manner. Land managers and resource specialists have many complex information requirements. They must integrate new sources of information into existing information systems, are confronted with determining the cost-effectiveness of all new techniques, and need useful products that provide adequate solutions to a wide range of problems. The results of remote sensing research, however, appear in technical reports and peer group journals which often are overly complex and replete with both necessary and unnecessary jargon. In addition, most previous research has ended by establishing the feasibility of a remote sensing procedure in a single, selected test area.

The gap between the user and the researcher of remote sensing technology can be bridged if the user is aware of what the technology has to offer, can determine that feasibility studies have been tested and are ready for operational use, has access to specialized equipment when needed, and is able to obtain hands-on training geared to specific needs. The remote sensing specialist, with a discipline science background, likewise must simplify, translate, and disseminate the offerings of the technology in a manner that is credible, understandable, and clearly illustrates both advantages and limitations. He also must test the technology to develop operational procedures that are ready for the manager.

The Geological Survey's EROS Data Center provides training and assistance in remote sensing techniques and thus assists in the technology transfer of remote sensing applications to users. The program includes conducting training courses and workshops, preparing and disseminating graphic and media educational aids, acting as a source of information for users interested in remote sensing technology, conducting demonstration projects in cooperation with Federal and State agencies, and demonstrating the application of automatic and semiautomatic image analysis systems to the solution of earth resource problems. The goals of this effort are to (1) see that operational remote sensing programs established in the Federal and State agencies have responsibility for inventorying, managing, and preserving our resources and environment, and (2) assist these agencies in developing new operational procedures that will upgrade their capability to carry out their responsibilities.

BUILDING THE BRIDGE TO THE USER
OR
WHY DON'T THEY USE OUR GOOD STUFF?

Fred C. Billingsley

Jet Propulsion Laboratory
Pasadena, California

Let us start by considering the statement that is usually made: "How can we give them the technology they need?" The setting of the statement (we...they) immediately sets up a dichotomy between us and them. In the context of the use of remotely sensed data for decision making by others than those who gathered it, the dichotomy is undoubtedly there. We (the technologists) have a set of technology in gathering and processing certain types of data that we think they (the decision makers) could use in their tasks. Thus, our view of the world might be as sketched in Fig. 1.
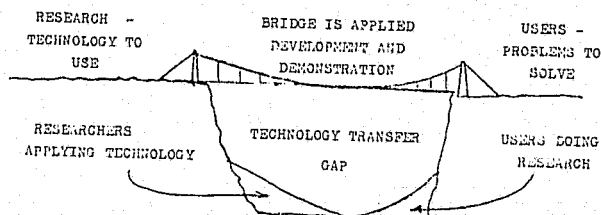

Fig. 1. The Technology Transfer Gap

We, the researchers, believe that we have developed a set of technology which we (on one cliff) are sure that they (on the other cliff) can use. Hopefully, if we can come to understand their tasks through mutual consultation, we may be able to suggest how our technology could be applied and thus bridge the gap. The gap is being filled to some extent on one side by researchers trying to apply their technology and on the other by some users doing some research, but this is a slow process; we feel the need to accelerate the process.

We cannot hope for this to occur by the users filling in their side by doing more research since they, by and large, are attempting to get their own tasks done with finite budgets and finite manpower. They are accomplishing these tasks with greater or less success and cannot afford to scrap what they are doing nor the data they are using in favor of a vague promise of improvement, nor even divert much of their resources toward researching the new methods. Thus, although some users are in a position to investigate the new methods, a general impasse exists.

Therefore, if the technology is to be used as fast as the technologists wish, we, the technologists, must build the bridge by:

● learning what data will be useful

● being sure it can be delivered before promising it

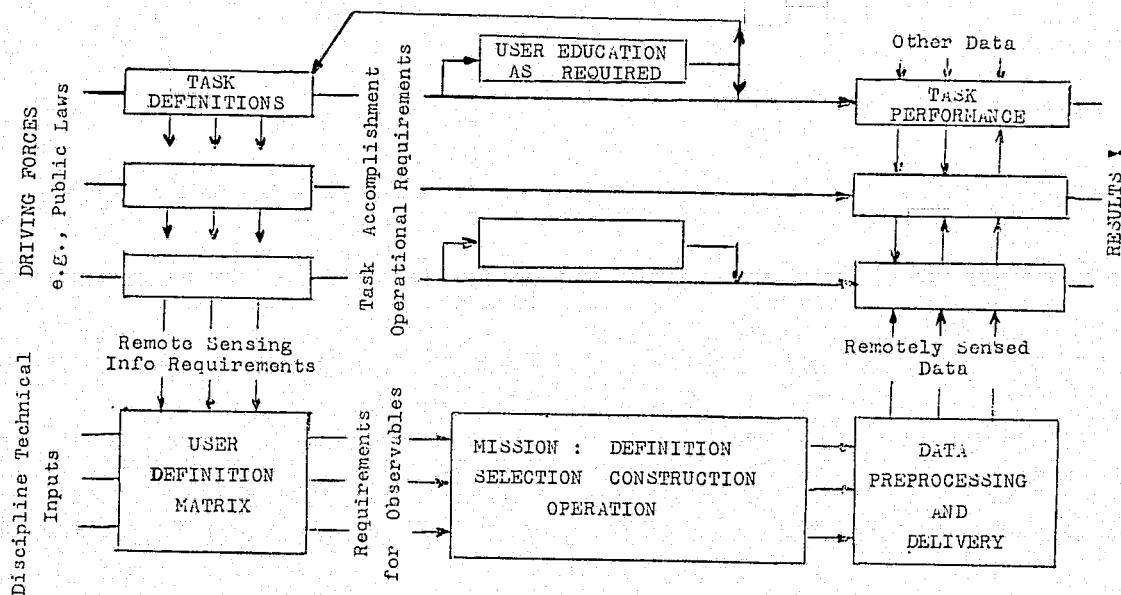● assuring that the potential users will know how to use it when it arrives

Fig. 2. Building the Remote Sensing Bridge

- delivering the data as promised, in suitable format

The flow diagram shows the general process in a somewhat simplified manner. The main path is across the top, from the (typical) public law driving forces to eventual successful results. This is the user's path — his task definitions will result in requirements for remotely sensed information and in a set of operational requirements/procedures.

A place for bridge building exists in the opportunity to demonstrate the available technology to the prospective user, so that by this education he is better able to define his task procedures and perhaps modify his original task definitions. Of course, this requires that the technologists properly understand the user's tasks in order to demonstrate something useful.

The task of the remote sensing community is to supply the users with the remotely sensed data. This bridge-building point has two facets: proper identification of the data which will satisfy the user's information needs, and then supplying the data. The first of these must be a mutual learning process in which the data technologist learns what the data user is trying to accomplish, works with him in the definition of useful, practical data, steers him away from technically unreasonable requests, and together with him defines the data analysis. This is a critical part of the process which must avoid the promise of data which cannot be delivered in the required time or the promise of results of analysis which cannot be accomplished.

It will often be found that at this point the data technologists and the potential users are speaking two different languages. It has been found useful to draw the two viewpoints together by working out tables of uses somewhat as follows:

MANAGEMENT NEED

Map the urban-suburban boundary of the Standard Metropolitan Statistical Areas.

INFORMATION REQUIRED

Population density to allow drawing the boundary at 1000 persons per square mile.

OBSERVABLE EFFECTS

The effect seems to be primarily one caused by variation in the (tree,grass)/(roof,pavement) ratio.

REMOTELY SENSED DATA REQUIRED

Spectral measurements in bands suitable for vegetation mapping: 0.4, 0.55, 0.7 μm for veg stress; thermal (10 μm) may help in the city.

Resolution element size: 80M adequate, 30M better.

Seasonal coverage to give added discrimination.

Geographical reference critical; will be registered to maps.

Data to be digital; will be incorporated into data base with other data on a geographical basis.

Two week maximum delay time to allow checking of critical areas against ground truth before any changes occur.

By the time such tables are complete, the user will understand what the technologist has to offer, and the technologist will have a group of requirements with which to synthesize missions.

Finally, there are the tasks of gathering and supplying the data. If the methodology outlined here is being followed, this is the point at which the spacecraft, sensor, and mission operations concepts are brought together, and prototype missions proposed. It will probably be found, in comparing prototype missions for several tiers of users, that there are overlapping requirements and that by intelligent regrouping of components optimum use of a limited set of missions will result.

In this mission set, it will probably be found that no one mission satisfies all the requirements for any given user, not will one mission be used exclusively by one user. The data from a given mission, therefore, must be prepared in such a manner that it can find the widest use by diverse users with a minimum of further specific processing (especially processing which might be duplicated by many users) required by each user before the data can be used.

Thus we are led to the final section of the bridge to the users: preprocessing the data to put it into the most suitable form for the widest range of users and delivering it in timely fashion. Consideration must be given at this point to the utilization of data from several missions by any given user, and therefore to the desirability of presenting it in such a way that data from several mission types can easily be interfaced.

From the users' viewpoint, it cannot be overemphasized that optimum use of the data is obtained only if the data is properly calibrated and presented in suitable format. For remotely sensed image data, this format requires radiometric and geometric rectifications, geographical referencing, and a data form (i.e., digital tapes and/or visible images) suitable to the analysis being performed. For some users of the data, very rapid delivery of the data is a requisite; the data system designers must attempt to accommodate these or at least specifically recognize what has been designed out.

Just as a physical bridge is not complete unless all sections are in place, so this hypothetical bridge to the users is not complete without all its sections. These

sections — yser/technologist mutual task definition, user procedure education, data gathering, and data preprocessing and timely delivery — must all be present in some form. Omission or slighting of any one will reduce traffic across the bridge, perhaps to the point of none at all, with the user doing his thing the old way, without using the remotely sensed data at all.

BRIDGE BUILDING

NASA would like to bridge the mutual task definition gap by establishing user needs committees covering subjects of national importance. They would cover such areas (but not necessarily be grouped) as renewable resources (for example, timber and agriculture), non-renewable resources, land use, coastal zone management, water resources, environmental monitoring, and navigation and communication. The commitees would consist of people from all layers of government and the private sector and cover the required disciplines. By this comprehensive cross section and by membership rotation, NASA can obtain continued inputs on the applications needs of the users.

Response to these expressed needs occurs when these needs are cross-compared, suitable observables are defined with required parameters, and missions synthesized and operated which take maximum advantage of the technology and provide the best available combination of mission parameters. Combined with well planned data preprocessing after receipt and timely data delivery, this will assure wide utilization of the data by a broad group of users.

To bridge the gap of user education and technology demonstration, NASA now has operating a group of Applications System Verification Tests (ASVTs). An ASVT, by definition, will demonstrate the total technical capability that is necessary to accomplish an applications objective. As the test develops, it is expected that the user commitment of resources will increase and NASA's will decrease. At the completion of ASVT activity, NASA will step out and the user will be self-reliant, continuing the activity on his own.

The final bridge, technical demonstration, is being considered by NASA in the form of regional centers. At such a center, a potential user could find out about the applications technology and obtain hands-on experience with some of the hardware. Staff would be available to provide assistance in pilot applications projects. The intent here would be to develop the understanding of the usefulness of the technology by the user, so that he will be able to apply it on his own. Until the user takes this last step, true technology transfer has not occurred. Therefore, both the ASVTs and the activities of the transfer centers will be monitored closely to assure that the projects are worthy, that the user really intends to take over at the end of the project, and that the projects are of wide enough interest and are well enough executed that there will be some horizontal diffusion to other users rather than be only a one-time application.

---

*On assignment to NASA Headquarters, Washington, DC.

THE NATIONAL TECHNICAL INFORMATION SERVICE

The National Technical Information Service of the U.S. Department of Commerce is the central source for the public sale of Government-sponsored research, development and engineering reports and other analyses prepared by Federal agencies, their contractors or grantees, or by Special Technology Groups. NTIS also is a central source for Federally generated machine processable data files.

NTIS ships about 20,000 information products daily as one of the world's leading processors of specialty information. It supplies its customers with about four million documents and microforms annually. The NTIS information collection exceeds 900,000 titles and all are available for sale. About 105,000 titles are stocked in multiple copies. Current lists of best selling reports describe those most in demand.

Customers may quickly locate summaries of interest from among some 480,000 Federally sponsored research reports completed and published from 1964 to date by using the agency's on-line computer search service (NTIS Search) or by referring to more than 1000 published Searches. About 60,000 new summaries and reports are added annually. An additional 180,000 description of ongoing research compiled by the Smithsonian Science Information Exchange, also are computer retrieval. Copies of the whole research reports on which the summaries are based, are sold by NTIS in paper or microfiche.

The NTIS Bibliographic Data File (on magnetic tape) includes unpublished research summaries and is available for lease. The computer products of other Federal agencies also are sold or leased by NTIS.

Current summaries of new research reports and other specialized information in various categories of interest are published in some 26 weekly newsletters, (Weekly Government Abstracts) and these are indexed. An all-inclusive bi-weekly journal (Government Reports Announcements & Index) is published for librarians, technical information specialists and those requiring all the summaries in a single volume.

A standing order microfiche service (SRIM) automatically provides subscribers with the full texts of research reports specially selected to satisfy individual requirements. Automatic distribution of paper copies also is available.

The foregoing and many more information products and services are described in the free NTIS general catalog (PR-154).

Additional services, such as the coordination, packaging and marketing of unusual information for individuals and organizations, may be specially designed.

The NTIS goal is to give timely information when it is needed. Should you have a problem or a request which requires fast action, telephone the Springfield Operations Center between 7:45 A.M. and 4:15 P.M., Monday thru Friday, or at the Washington Information Center and Bookstore between 9 A.M. and 4:30 P.M. Telex and Telecopier service operates 24 hours, seven days.

To Check On Orders
(703) 557-4660

Subscriptions
(703) 557-4630

Computer Products
(703) 557-4763

New Orders for Documents & Reports
(703) 557-4650

NTISearches
(703) 557-4640

SRIM
(703) 557-4630

Pickup Orders
Washington Information Center & Bookstore
between 9 A.M. and 4:30 P.M.
(202) 724-3382

At Springfield Operations Center
(703) 557-4604

Accounting
(703) 557-4628

Rush Order Service
(703) 557-4700

Other Information
(703) 557-4600

Telecopier or 3M Facsimile Service
(703) 321-8547

Telex-89-9405